

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:11:13 ; Search time 126 Seconds  
(without alignments)  
1004.613 Million cell updates/sec

Title: US-10-650-369-22  
Perfect score: 2278  
Sequence: 1 MKKITGIILLLLAVIILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		SUMMARIES			
Result No.	Score	Query		DB ID	Description
		Match	Length		
1	2278	100.0	448	5	AAM50664 Streptoco
2	1656.5	72.7	336	5	AAM50665 Streptoco
3	1656.5	72.7	336	5	AAM50639 Streptoco
4	1655.5	72.7	336	5	ABP29960 Streptoco
5	1655.5	72.7	336	6	ABU46455 Protein e
6	1655.5	72.7	345	5	ABP29106 Streptoco
7	1652.5	72.5	336	2	AAR56486 Plasmin r
8	1652.5	72.5	336	4	AAV85681 Streptoco
9	1564.5	68.7	336	5	ABP30758 Streptoco
10	1559.5	68.5	336	5	AAM50667 Streptoco
11	1559.5	68.5	336	5	AAM50641 Streptoco
12	1557.5	68.4	336	5	AAM50666 Streptoco
13	1557.5	68.4	336	5	AAM50640 Streptoco
14	1535	67.4	335	6	ABU02516 S. pneumo
15	1535	67.4	359	4	AAU37576 Streptoco
16	1535	67.4	359	6	ABU46262 Protein e
17	1534.5	67.4	336	5	AAM50669 Streptoco
18	1534.5	67.4	336	5	AAM50643 Streptoco
19	1534	67.3	359	4	AAU38000 Streptoco
20	1530.5	67.2	336	5	AAM50668 Streptoco
21	1530.5	67.2	336	5	AAM50642 Streptoco
22	1521	66.8	337	6	ABU44330 Protein e
23	1491	65.5	333	2	AAW55089 Streptoco
24	1491	65.5	333	5	ABP54583 S. pneumo
25	1491	65.5	333	7	ADC45135 S. pneumo

26	1357.5	59.6	336	5	ABB55601 Lactococc
27	1282	56.3	333	4	AAU35254 Enterococ
28	1282	56.3	333	6	ABU29370 Protein e
29	1275	56.0	336	5	ABB53868 Lactococc
30	1270	55.8	333	7	ADC95142 E. faeciu
31	1224.5	53.8	335	6	ABU25342 Protein e
32	1205.5	52.9	336	5	ABB48810 Listeria
33	1205.5	52.9	336	6	ABU32568 Protein e
34	1203	52.8	334	6	ABU23595 Protein e
35	1166	51.2	335	6	ABU43810 Protein e
36	1145	50.3	334	6	ABU37737 Protein e
37	1141	50.1	357	6	ABP81005 N. gonorr
38	1141	50.1	357	6	ABU37077 Protein e
39	1138.5	50.0	336	4	AAG83058 S. epider
40	1138.5	50.0	336	6	ABU42966 Protein e
41	1138.5	50.0	348	5	ABP40668 Staphyloc
42	1132.5	49.7	336	4	AAU36947 Staphyloc
43	1132.5	49.7	336	4	AAU34122 Staphyloc
44	1132.5	49.7	336	6	ABJ19195 Pathogen
45	1132.5	49.7	336	6	ABU16073 Protein e

ALIGNMENTS

RESULT 1  
AAM50664  
ID AAM50664 standard; protein; 448 AA.  
XX  
AC AAM50664;  
XX  
DT 29-AUG-2003 (revised)  
DT 08-APR-2002 (first entry)  
XX  
DE Streptococcus GapC multiepitope fusion PolyGap4.  
XX  
KW PolyGap4; GapC; plasmin binding protein; epitope; infection; vaccine; immunisation; mastitis; therapy.  
KW  
XX Streptococcus dysgalactiae.  
OS Streptococcus agalactiae.  
OS Streptococcus parauberis.  
OS Chimeric.  
XX  
PN WO200196379-A2.  
XX  
PD 20-DEC-2001.  
XX

PF 11-JUN-2001; 2001WO-CA000836.  
XX  
PR 12-JUN-2000; 2000US-0211247P.  
XX  
PA (UYSA-) UNIV SASKATCHEWAN.  
XX  
PI Potter AA, Perez-Casal J, Fontaine M;  
XX WPI; 2002-098051/13.  
DR N-PSDB; ABA91327.  
XX  
PT Novel GapC multiple epitope fusion polypeptide comprising antigenic determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S. parauberis, or S. iniae GapC protein useful for treating mastitis in vertebrates.  
XX  
PS Claim 9; Fig 6A-C; 116pp; English.  
XX  
CC The present sequence is that of a novel multiple epitope fusion protein, designated PolyGap4, comprising the entire amino acid sequence of the Streptococcus dysgalactiae GapC plasmin binding protein in addition to unique amino acid sequences from the Streptococcus parauberis and Streptococcus agalactiae GapC proteins. The multiple epitope protein is produced in host cells transformed with an expression vector comprising a chimeric gene (see ABA91327) encoding the protein. PolyGal4 is an example

CC of novel GapC multiple epitope fusion proteins of the invention that  
CC comprise epitopes from 1 or more of *S. dysgalactiae*, *S. agalactiae*, *S.*  
CC *parauberis*, *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-  
CC 69). The multiple epitope fusion proteins are used in claimed vaccines  
CC for treating or preventing a bacterial infection in a vertebrate,  
CC especially a streptococcal infection, and particularly mastitis. They are  
CC also used in claimed methods of detecting *Streptococcus* antibodies. The  
CC multiple epitope protein is capable of eliciting broad immunity against a  
CC variety of streptococcal infections while minimising the number of  
CC antigens present in the final formulation and concomitantly reducing  
CC production costs. (Updated on 29-AUG-2003 to standardise OS field)  
XX  
SQ Sequence 448 AA;

Query Match 100.0%; Score 2278; DB 5; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.8e-180;  
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKITGIILLAVIILSACQANYGSMVVKVGINGFGRIGRLAFRIQNVGVEVTRIN 60  
Db 1 MKKITGIILLAVIILSACQANYGSMVVKVGINGFGRIGRLAFRIQNVGVEVTRIN 60  
QY 61 DLTDPNMLAHLKDYDTTQGRFDGTVEVKEGGFEVGNFIKVSARDPENIDWATDGEIV 120  
Db 61 DLTDPNMLAHLKDYDTTQGRFDGTVEVKEGGFEVGNFIKVSARDPENIDWATDGEIV 120  
QY 121 LEALEGTVEVKDGGFDVNGKFIKVSAREKDPQIDWATDGEIVLEIDGTVEVKEGGFEVN 180  
Db 121 LEALEGTVEVKDGGFDVNGKFIKVSAREKDPQIDWATDGEIVLEIDGTVEVKEGGFEVN 180  
QY 181 GQFVKVSAREPANIDWATDGEIVLEATSFPAKKEAAEKHLHANGAKKVITAPGGNDV 240  
Db 181 GQFVKVSAREPANIDWATDGEIVLEATSFPAKKEAAEKHLHANGAKKVITAPGGNDV 240  
QY 241 KTVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHAYTGDQMIL 300  
Db 241 KTVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHAYTGDQMIL 300  
QY 301 DGPHRGDLRRARAGAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRVPVPTGSVTELVV 360  
Db 301 DGPHRGDLRRARAGAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRVPVPTGSVTELVV 360  
QY 361 TLDKNVSVDEINAAKKAASNDISFGYTEDPIVSSDIVGSYGLFDATQTKVMEVDGSQLV 420  
Db 361 TLDKNVSVDEINAAKKAASNDISFGYTEDPIVSSDIVGSYGLFDATQTKVMEVDGSQLV 420  
QY 421 KVVSVYDNEMSYTAQLVRTLEYFAKIAK 448  
Db 421 KVVSVYDNEMSYTAQLVRTLEYFAKIAK 448

RESULT 2  
AAM50665  
ID AAM50665 standard; protein; 336 AA.  
XX  
AC AAM50665;  
XX  
DT 08-APR-2002 (first entry)  
XX  
DE Streptococcus dysgalactiae gapC plasmin binding protein DysGapC.  
XX  
KW DysGapC; GapC; plasmin binding protein; epitope; infection; vaccine;  
XX immunisation; mastitis; therapy.  
OS Streptococcus dysgalactiae.  
XX  
PN WO200196379-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 11-JUN-2001; 2001WO-CA000836.  
XX  
PR 12-JUN-2000; 2000US-0211247P.

XX (UYSA-) UNIV SASKATCHEWAN.  
PA Potter AA, Perez-Casal J, Fontaine M;  
XX  
PI WPI; 2002-098051/13.  
DR N-PSDB; ABA91328.  
DR  
XX  
PT Novel GapC multiple epitope fusion polypeptide comprising antigenic  
PT determinant of *Streptococcus dysgalactiae*, *S. agalactiae*, *S. uberis*, *S.*  
PT *parauberis*, or *S. iniae* GapC protein useful for treating mastitis in  
PT vertebrates.  
XX  
PS Claim 8; Fig 1A-B; 116pp; English.  
XX  
CC The present sequence is that of the GapC plasmin binding protein,  
CC DysGapC, of *Streptococcus dysgalactiae* ATCC 43078, an isolate from a case  
CC of bovine mastitis. The invention relates to novel GapC multiple epitope  
CC fusion proteins that comprise epitopes from 1 or more of *Streptococcus*  
CC *dysgalactiae*, *Streptococcus agalactiae*, *Streptococcus parauberis*,  
CC *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-69). A claimed  
CC example is PolyGap4 (see AAM50664). Expression vectors and host cells for  
CC production of the multiple epitope fusion proteins are provided. The  
CC multiple epitope proteins are used in claimed vaccines for treating or  
CC preventing a bacterial infection in a vertebrate, especially a  
CC streptococcal infection, and particularly mastitis. They are also used in  
CC claimed methods of detecting *Streptococcus* antibodies. The multiple  
CC epitope proteins are capable of eliciting broad immunity against a  
CC variety of streptococcal infections while minimising the number of  
CC antigens present in the final formulation and concomitantly reducing  
CC production costs  
XX  
SQ Sequence 336 AA;

Query Match 72.7%; Score 1656.5; DB 5; Length 336;  
Best Local Similarity 79.6%; Pred. No. 4.9e-129;  
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;  
QY 28 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 87  
Db 1 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 60  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAR 147  
Db 61 KEGGFEVNGNFIKVSARDPE----- 81  
QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIDWATDGEIVLE 207  
Db 82 -----NIDWATDGEIVLE 95  
QY 208 ATSPFAKKEAAEKHLHANGAKKVITAPGGNDVKTVFNTNHDILDGTETVISGASCTTN 267  
Db 96 ATGFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVFNTNHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMI LDGPHRGDLRRARAGAANIVPNSTGAA 327  
Db 156 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMI LDGPHRGDLRRARAGAANIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVTLDKNVSVDEINAAKKAASNDISFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVTLDKNVSVDEINAAKKAASNDISFGYTE 275  
QY 388 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVVSVYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVVSVYDNEMSYTAQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336

RESULT 3  
AAM50639

ID AAM50639 standard; protein; 336 AA.  
XX  
AC AAM50639;  
XX  
DT 04-APR-2002 (first entry)  
XX  
DE Streptococcus dysgalactiae gapC plasmin binding protein.  
XX  
KW GapC; plasmin-binding protein; DysgalGapC; infection; mastitis; vaccine;  
KW diagnosis; therapy.  
XX  
OS Streptococcus dysgalactiae.  
XX  
PN WO200196381-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 11-JUN-2001; 2001WO-CA000838.  
XX  
PR 12-JUN-2000; 2000US-0211022P.  
XX  
PA (UYSA-) UNIV SASKATCHEWAN.  
XX  
PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;  
XX  
XX WPI; 2002-130725/17.  
DR N-PSDB; ABA91248.  
XX  
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,  
PT S.uberis, S.parauberis, or S.iniae, useful as vaccine component for  
PT treating streptococcal infection which causes mastitis in vertebrates.  
XX  
PS Claim 1(a); Fig 1A-B; 107pp; English.  
XX  
XX The present sequence is that of the GapC plasmin binding protein  
CC (DysgalGapC) of Streptococcus dysgalactiae ATCC 43078, a clinical isolate  
CC from bovine mastitis. It is encoded by the GapC gene given in ABA91248.  
CC  
CC GapC protein, which has no signal sequence or membrane anchor domain, is  
CC capable of eliciting an immune response in a vertebrate. The invention  
CC provides the GapC genes and proteins of 5 Streptococcus species, as well  
CC as recombinant vectors, host cells and vaccine compositions comprising  
CC GapC polynucleotides or proteins. The vaccines are used to treat or  
CC prevent a bacterial infection, especially a streptococcal infection, and  
CC mastitis in particular (claimed). GapC proteins are also used in claimed  
CC methods for detecting GapC antibodies, and to raise antibodies that are  
CC used in claimed methods for detecting GapC proteins  
XX  
SQ Sequence 336 AA;

Query Match 72.7%; Score 1656.5; DB 5; Length 336;  
Best Local Similarity 79.6%; Pred. No. 4.9e-129;  
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVETRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 87  
|||  
DB 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVETRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60  
|||  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVS AE 147  
|||  
DB 61 KEGGFEVNGNFIKVSARDPE----- 81  
|||  
QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207  
|||  
DB 82 -----NIDWATDGV EIVLE 95  
|||  
QY 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 267  
|||  
DB 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 155  
|||  
QY 268 CLAPMAKALHDAPGIQKGLMTTIHAYTGQMLDGP HRRGDLRRARAGAAANIVPNSTGAA 327  
|||  
DB 156 CLAPMAKALHDAPGIQKGLMTTIHAYTGQMLDGP HRRGDLRRARAGAAANIVPNSTGAA 215  
|||

QY 328 KAIGLVIPELNGKLDGAAQRPVPPTGVS TELVVTLDKNVSVDEINAAAMKAASNDSPGYTE 387  
|||  
DB 216 KAIGLVIPELNGKLDGAAQRPVPPTGVS TELVVTLDKNVSVDEINAAAMKAASNDSPGYTE 275  
|||  
QY 388 DPIVSSDIVGVSIGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTL EYFAKIA 447  
|||  
DB 276 DPIVSSDIVGVSIGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTL EYFAKIA 335  
|||  
QY 448 K 448  
DB 336 K 336  
|||  
RESULT 4  
ABP29960  
ID ABP29960 standard; protein; 336 AA.  
XX  
AC ABP29960;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Streptococcus polypeptide SEQ ID NO 9096.  
XX  
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
XX  
OS Streptococcus pyogenes.  
XX  
PN WO200234771-A2.  
XX  
PD 02-MAY-2002.  
XX  
PS 29-OCT-2001; 2001WO-GB004789.  
PF  
XX 27-OCT-2000; 2000GB-00026333.  
PR 24-NOV-2000; 2000GB-00028727.  
PR 07-MAR-2001; 2001GB-00005640.  
XX  
XX (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;  
PI Tettelin H;  
XX  
XX WPI; 2002-352536/38.  
DR N-PSDB; ABN70591.  
XX  
XX New Streptococcus protein for the treatment or prevention of infection or  
PT disease caused by Streptococcus bacteria, such as meningitis, and for  
PT detecting a compound that binds to the protein.  
XX  
PS Claim 1; Page 4031; 4525pp; English.  
XX  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
XX Streptococcus proteins  
SQ Sequence 336 AA;



Query Match		72.7%;	Score 1655.5;	DB 5;	Length 336;
Best Local Similarity		79.3%;	Pred. No. 5.9e-129;		
Matches 334;		Conservative 1;	Mismatches 1;	Indels 85;	Gaps 1;
QY	28	MVVKVGINGFGRIGRLAFRRRIQNV	EGVEVTRINDLTDPNMLAHLKYD	TQGRFDGTVEV	87
Db	1	MVVKVGINGFGRIGRLAFRRRIQNV	EGVEVTRINDLTDPNMLAHLKYD	TQGRFDGTVEV	60
QY	88	KEGGFEVNGNFIKVS	AERDPENIDWATDGV	EVLEALEGTVEVKDGGFDVNGKFIKVS	AE 147
Db	61	KEGGFEVNGNFIKVS	AERDPE-----	-----	81
QY	148	KDPEQIDWATDGV	EVLEIDGTVEVKEGGFEVNGQFVKV	SAEREPANIDWATDGV	EVLE 207
Db	82	-----	-----	-----NIDWATDGV	EVLE 95
QY	208	ATSFPAKKEAAEKHLHANGAKKVVIT	APGGNDVKTVVNTNHDILDGTETV	ISGASCTTN	267
Db	96	ATGFFAKKEAAEKHLHANGAKKVVIT	APGGNDVKTVVNTNHDILDGTETV	ISGASCTTN	155
QY	268	CLAPMAKALHDAFGIQKGLMTTTHAY	TGDMILDGPHRGDLRRARAGAA	NI	VPNSTGAA 327
Db	156	CLAPMAKALHDAFGIQKGLMTTTHAY	TGDMILDGPHRGDLRRARAGAA	NI	VPNSTGAA 215
QY	328	KAIGLVIPELNGKLDGAAQRPVPTGS	VTSELVVTLDKNVSVDEINAAK	KAASND	SFGYTE 387
Db	216	KAIGLVIPELNGKLDGAAQRPVPTGS	VTSELVVTLDKNVSVDEINAAK	KAASND	SFGYTE 275
QY	388	DPIVSSDIVGSYGS	LFDTATQTKVMEVDGSQLVKV	SVWYDNEMSYTAQLV	RTLEYFAKIA 447
Db	276	DPIVSSDIVGSYGS	LFDTATQTKVMEVDGSQLVKV	SVWYDNEMSYTAQLV	RTLEYFAKIA 335
QY	448	K 448			
Db	336	K 336			

RESULT 5	
ABU46455	
ID	ABU46455 standard; protein; 336 AA.
XX	
AC	ABU46455;
XX	
DT	19-JUN-2003 (first entry)
XX	
DE	Protein encoded by Prokaryotic essential gene #31982.
XX	
KW	Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	
OS	Streptococcus pyogenes.
XX	
PN	WO200277183-A2.
XX	
PD	03-OCT-2002.
XX	
PF	21-MAR-2002; 2002WO-US009107.
XX	
PR	21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-0362699P.
XX	
PA	(ELIT-) ELITRA PHARM INC.
XX	
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	
DR	WPI; 2003-029926/02.
DR	N-PSDB; ACA50325.
XX	
PT	New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 74379; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

Sequence 336 AA;

Query Match		72.7%;	Score 1655.5;	DB 6;	Length 336;
Best Local Similarity		79.3%;	Pred. No. 5.9e-129;		
Matches 334;		Conservative 1;	Mismatches 1;	Indels 85;	Gaps 1;
QY	28	MVVKVGINGFGRIGRLAFRRRIQNV	EGVEVTRINDLTDPNMLAHLKYD	TQGRFDGTVEV	87
Db	1	MVVKVGINGFGRIGRLAFRRRIQNV	EGVEVTRINDLTDPNMLAHLKYD	TQGRFDGTVEV	60
QY	88	KEGGFEVNGNFIKVS	AERDPENIDWATDGV	EVLEALEGTVEVKDGGFDVNGKFIKVS	AE 147
Db	61	KEGGFEVNGNFIKVS	AERDPE-----	-----	81
QY	148	KDPEQIDWATDGV	EVLEIDGTVEVKEGGFEVNGQFVKV	SAEREPANIDWATDGV	EVLE 207
Db	82	-----	-----	-----NIDWATDGV	EVLE 95
QY	208	ATSFPAKKEAAEKHLHANGAKKVVIT	APGGNDVKTVVNTNHDILDGTETV	ISGASCTTN	267
Db	96	ATGFFAKKEAAEKHLHANGAKKVVIT	APGGNDVKTVVNTNHDILDGTETV	ISGASCTTN	155
QY	268	CLAPMAKALHDAFGIQKGLMTTTHAY	TGDMILDGPHRGDLRRARAGAA	NI	VPNSTGAA 327
Db	156	CLAPMAKALHDAFGIQKGLMTTTHAY	TGDMILDGPHRGDLRRARAGAA	NI	VPNSTGAA 215
QY	328	KAIGLVIPELNGKLDGAAQRPVPTGS	VTSELVVTLDKNVSVDEINAAK	KAASND	SFGYTE 387
Db	216	KAIGLVIPELNGKLDGAAQRPVPTGS	VTSELVVTLDKNVSVDEINAAK	KAASND	SFGYTE 275
QY	388	DPIVSSDIVGSYGS	LFDTATQTKVMEVDGSQLVKV	SVWYDNEMSYTAQLV	RTLEYFAKIA 447
Db	276	DPIVSSDIVGSYGS	LFDTATQTKVMEVDGSQLVKV	SVWYDNEMSYTAQLV	RTLEYFAKIA 335
QY	448	K 448			



Db	336 K 336		
RESULT 6			
ABP29106			
ID	ABP29106 standard; protein; 345 AA.		
XX			
AC	ABP29106;		
XX			
DT	02-JUL-2002 (first entry)		
XX			
DE	Streptococcus polypeptide SEQ ID NO 7388.		
XX			
KW	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;		
KW	group A streptococcus; Streptococcus pyogenes; antibacterial;		
KW	antiinflammatory; infection; vaccine; meningitis; gene therapy.		
XX			
OS	Streptococcus pyogenes.		
XX			
PN	WO200234771-A2.		
XX			
PD	02-MAY-2002.		
XX			
PF	29-OCT-2001; 2001WO-GB004789.		
XX			
PR	27-OCT-2000; 2000GB-00026333.		
PR	24-NOV-2000; 2000GB-00028727.		
PR	07-MAR-2001; 2001GB-00005640.		
XX			
PA	(CHIR-) CHIRON SPA.		
PA	(GENO-) INST GENOMIC RES.		
XX			
PI	Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;		
PI	Tettelin H;		
XX			
DR	WPI; 2002-352536/38.		
DR	N-PSDB; ABN69737.		
XX			
PT	New Streptococcus protein for the treatment or prevention of infection or		
PT	disease caused by Streptococcus bacteria, such as meningitis, and for		
PT	detecting a compound that binds to the protein.		
XX			
PS	Claim 1; Page 3888; 4525pp; English.		
XX			
CC	The invention relates to a protein (ABP25413-ABP30895) from group B		
CC	Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS		
CC	(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in		
CC	the specification. The proteins have antibacterial and antiinflammatory		
CC	activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and		
CC	antibodies that bind (I) are used in the manufacture of medicaments for		
CC	the treatment or prevention of infection or disease caused by		
CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.		
CC	Nucleic acids encoding (I) are used to detect Streptococcus in a		
CC	biological sample. (I) is used to determine whether a compound binds to		
CC	(I). A composition comprising (I) or a nucleic acid encoding (I), may be		
CC	used as a vaccine or diagnostic composition. The disease caused by		
CC	Streptococcus that is prevented or treated may be meningitis. Nucleic		
CC	acid encoding (I) may be used to recombinantly produce (I) and may be		
CC	used in gene therapy. Antibodies to (I) are used for affinity		
CC	chromatography, immunoassays, and distinguishing/identifying		
CC	Streptococcus proteins		
XX			
SQ	Sequence 345 AA;		
Query Match	72.7%; Score 1655.5; DB 5; Length 345;		
Best Local Similarity	79.3%; Pred. No. 6.1e-129;		
Matches 334; Conservative	1; Mismatches 1; Indels 85; Gaps 1;		
QY	28 MVVKVGINGFGRIGRLAFRRIQNVGEVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87		
Db	10 MVVKVGINGFGRIGRLAFRRIQNIEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 69		
QY	88 KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVDGGFDVNGKFIKVS AE 147		

Db	70	KEGGFEVNGNFIKVSARDPE-----	90
QY	148	KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIDWATDGVVEIVLE	207
Db	91	-----NIDWATDGVVEIVLE	104
QY	208	ATSEFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	267
Db	105	ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	164
QY	268	CLAPMAKALHDAFGIQKGLMTTHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA	327
Db	165	CLAPMAKALHDAFGIQKGLMTTHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA	224
QY	328	KAIGLVIPELNGKLDGAAQRVPVPTGVTTELVTLDKNVSVDEINAAMKAASNDSFGYTE	387
Db	225	KAIGLVIPELNGKLDGAAQRVPVPTGVTTELVTLDKNVSVDEINAAMKAASNDSFGYTE	284
QY	388	DPIVSSDIVGSYSGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA	447
Db	285	DPIVSSDIVGSYSGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA	344
QY	448	K 448	
Db	345	K 345	
RESULT 7			
AA	AAR56486		
ID	AAR56486 standard; protein; 336 AA.		
XX			
AC	AAR56486;		
XX			
DT	25-MAR-2003 (revised)		
DT	22-FEB-1995 (first entry)		
XX			
DE	Plasmin receptor.		
XX			
KW	Plasmin receptor; isolate 64/14; plasmin; alpha-2-antiplasmin;		
KW	plasminogen activator; bleeding; reocclusion; thrombosis;		
KW	pulmonary embolism; clots.		
XX			
OS	Streptococcus pyogenes.		
XX			
PN	US5328996-A.		
XX			
PD	12-JUL-1994.		
XX			
PF	10-AUG-1992; 92US-00928462.		
XX			
PR	29-MAR-1989; 89US-00330849.		
PR	16-MAY-1990; 90US-00524411.		
XX			
PA	(UYFL ) UNIV FLORIDA RES FOUND INC.		
XX			
PI	Von Mering G, Broder C, Boyle MDP, Lottenberg R;		
XX			
DR	WPI; 1994-225327/27.		
DR	N-PSDB; AAQ70705.		
XX			
PT	New DNA encoding bacterial plasmin receptor - useful as thrombolytic		
PT	agents, used with plasminogen activator or bound to plasmin, also useful		
PT	in vaccines against bacterial infection.		
XX			
PS	Claim 1; Col 27-30; 19pp; English.		
XX			
CC	This sequence represents the S. pyogenes plasmin receptor. The DNA		
CC	encoding this sequence was isolated from the S. pyogenes clinical isolate		
CC	64/14. The plasmin receptor has a very high affinity for plasmin which,		
CC	when bound, retains its enzymatic activity but is not regulated		
CC	(inhibited) by alpha-2-antiplasmin. The receptor protein, when		
CC	administered concurrently or sequentially, prolongs the activity of		

CC plasminogen activator (PA) so allows a reduction in dose, and thus lowers  
CC the risk of bleeding, and may prevent reocclusion of blood vessels. The  
CC protein may be coupled to a fibrin-specific monoclonal antibody to  
CC provide targeting to clots. The plasmin receptor may be useful in human  
CC or veterinary medicine, for treatment of thrombosis and pulmonary  
CC embolism, and for solubilising clots in catheters or shunts. (Updated on  
CC 25-MAR-2003 to correct PF field.)  
XX  
SQ Sequence 336 AA;

Query Match 72.5%; Score 1652.5; DB 2; Length 336;  
Best Local Similarity 79.1%; Pred. No. 1e-128;  
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;  
QY 28 M V V K V G I N G F G R I G R L A F R R I O N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 87  
Db 1 M V V K V G I N G F G R I G R L A F R R I O N I E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60  
QY 88 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F D V N G K F I K V S A E 147  
Db 61 K E G G F E V N G N F I K V S A E R D P E ----- 81  
QY 148 K D P E Q I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E 207  
Db 82 -----N I D W A T D G V E I V L E 95  
QY 208 A T S F F A K K E A A E K H L H A N G A K K V I T A P G G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 267  
Db 96 A T G F F A K K E A A E K H L H A N G A K K V I T A P G G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 155  
QY 268 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 327  
Db 156 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 215  
QY 328 K A I G L V I P E L N G K L D G A A Q R V P T G S V T E L V V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 387  
Db 216 K A I G L V I P E L N G K L D G A A Q R V P T G S V T E L V V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 275  
QY 388 D P I V S S D I V G S Y G S L F D A T Q T K V M E V D G S Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 447  
Db 276 D P I V S S D I V G S Y G S L F D A T Q T K V M E V D G S Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 335  
QY 448 K 448  
Db 336 K 336

RESULT 8  
AA Y85681  
ID AAY85681 standard; protein; 336 AA.  
XX  
AC AAY85681;

XX  
DT 13-FEB-2001 (first entry)  
XX  
DE Streptococcal plasmin receptor amino acid sequence.  
XX  
KW Plasmin receptor; plr; immune response; vaccination; antibacterial;  
KW Streptococcal infection; antihelminthic.  
XX

OS Streptococcus pyogenes.  
XX  
PN US6136323-A.  
XX  
PD 24-OCT-2000.  
XX  
PF 11-JUL-1994; 94US-00273247.  
XX  
PR 29-MAR-1989; 89US-00330849.  
PR 16-MAY-1990; 90US-00524411.  
PR 10-AUG-1992; 92US-00928462.  
XX

PA (UYFL ) UNIV FLORIDA RES FOUND INC.

XX  
PI  
XX  
DR WPI; 2001-006210/01.  
XX N-PSDB; AAC66140.  
PT Raising an immune response in a mammal, especially for vaccination  
PT against group A streptococcal infections, comprises administering an  
PT isolated streptococcal plasmin receptor protein.  
XX  
PS Claim 2; Col 27-30; 17pp; English.  
XX  
CC Polynucleotide sequence AAC66140 encodes the Streptococcal plasmin  
CC receptor protein AAY85681. The protein is used in a method for raising an  
CC immune response in a mammal. The method comprises administering the  
CC plasmin receptor protein. The method is useful as a vaccination against  
CC group A Streptococcal infections and potentially against a broad range of  
CC infections associated with pathogens expressing glyceraldehyde-3-  
CC phosphate dehydrogenase (GAPDH) or GAPDH-like proteins at their cell  
CC surfaces. The vaccine has antibacterial and antihelminthic activity  
XX  
SQ Sequence 336 AA;

Query Match 72.5%; Score 1652.5; DB 4; Length 336;  
Best Local Similarity 79.1%; Pred. No. 1e-128;  
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;  
QY 28 M V V K V G I N G F G R I G R L A F R R I O N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 87  
Db 1 M V V K V G I N G F G R I G R L A F R R I O N I E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60  
QY 88 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F D V N G K F I K V S A E 147  
Db 61 K E G G F E V N G N F I K V S A E R D P E ----- 81  
QY 148 K D P E Q I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E 207  
Db 82 -----N I D W A T D G V E I V L E 95  
QY 208 A T S F F A K K E A A E K H L H A N G A K K V I T A P G G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 267  
Db 96 A T G F F A K K E A A E K H L H A N G A K K V I T A P G G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 155  
QY 268 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 327  
Db 156 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 215  
QY 328 K A I G L V I P E L N G K L D G A A Q R V P T G S V T E L V V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 387  
Db 216 K A I G L V I P E L N G K L D G A A Q R V P T G S V T E L V V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 275  
QY 388 D P I V S S D I V G S Y G S L F D A T Q T K V M E V D G S Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 447  
Db 276 D P I V S S D I V G S Y G S L F D A T Q T K V M E V D G S Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 335  
QY 448 K 448  
Db 336 K 336

RESULT 9  
ABP30758  
ID ABP30758 standard; protein; 336 AA.  
XX  
AC ABP30758;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Streptococcus polypeptide SEQ ID NO 10692.  
XX  
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW Group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB004789.

XX PR 27-OCT-2000; 2000GB-00026333.

XX PR 24-NOV-2000; 2000GB-00028727.

XX PR 07-MAR-2001; 2001GB-00005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;

XX PI Tettelin H;

XX PR WPI; 2002-352536/38.

XX DR N-PSDB; ABN71389.

XX PT New Streptococcus protein for the treatment or prevention of infection or

XX PT disease caused by Streptococcus bacteria, such as meningitis, and for

XX PT detecting a compound that binds to the protein.

XX PS Claim 1; Page 4179; 4525pp; English.

XX CC The invention relates to a protein (ABP25413-ABP30895) from group B

CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

CC the specification. The proteins have antibacterial and antiinflammatory

CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and

CC antibodies that bind (I) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to

CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

XX CC Streptococcus proteins

SQ Sequence 336 AA;

Query Match 68.7%; Score 1564.5; DB 5; Length 336;

Best Local Similarity 74.1%; Pred. No. 2.1e-121;

Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;

QY 28 MVVKVGINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTV 87

DB 1 MVVKVGINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRF----- 54

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGV EIVLEALEGTV EVDGDFVNGKFIKVS 147

DB 55 ----- 54

QY 148 KDEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207

DB 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 267

DB 96 ATGFFASKEAEQIHENGAKKVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGP HRRGDLRRARAGAAIVPNSTGAA 327

DB 156 CLAPMAKALQDNFGVKQGLMTTHAYTGDQMLDGP HRRGDLRRARAGAAIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLDKNVSVDEINAAKAAANDSFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLEKDVVEEVNAAMKAAANDSYGYTE 275

QY 388 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 447

Db 276 DPIVSSDIVGSYGSFLFDATQTKVQTVDGNQLVKVSWYDNMSYTSQLVRLTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 10

AAM50667

ID AAM50667 standard; protein; 336 AA.

XX AAM50667;

AC AAM50667;

XX DT 08-APR-2002 (first entry)

XX DE Streptococcus uberis gapC plasmin binding protein UberGapC.

XX KW UberGapC; GapC; plasmin binding protein; epitope; infection; vaccine;

XX KW immunisation; mastitis; therapy.

XX OS Streptococcus uberis.

XX PN WO200196379-A2.

XX PD 20-DEC-2001.

XX PF 11-JUN-2001; 2001WO-CA000836.

XX PR 12-JUN-2000; 2000US-0211247P.

XX PA (UUSA-) UNIV SASKATCHEWAN.

XX PI Potter AA, Perez-Casal J, Fontaine M;

XX DR WPI; 2002-098051/13.

XX DR N-PSDB; ABA91330.

XX PT Novel GapC multiple epitope fusion polypeptide comprising antigenic

PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.

PT parauberis, or S. iniae GapC protein useful for treating mastitis in

PT vertebrates.

XX PS Claim 8; Fig 3A-B; 116pp; English.

XX CC The present sequence is that of the GapC plasmin binding protein,

CC UberGapC, of Streptococcus uberis ATCC 9927. The invention relates to

CC novel GapC multiple epitope fusion proteins that comprise epitopes from 1

CC or more of Streptococcus dysgalactiae, Streptococcus agalactiae,

CC Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae

CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).

CC Expression vectors and host cells for production of the multiple epitope

CC fusion proteins are provided. The multiple epitope proteins are used in

CC claimed vaccines for treating or preventing a bacterial infection in a

CC vertebrate, especially a streptococcal infection, and particularly

CC mastitis. They are also used in claimed methods of detecting

CC Streptococcus antibodies. The multiple epitope proteins are capable of

CC eliciting broad immunity against a variety of streptococcal infections

CC while minimising the number of antigens present in the final formulation

CC and concomitantly reducing production costs

XX SQ Sequence 336 AA;

Query Match 68.5%; Score 1559.5; DB 5; Length 336;

Best Local Similarity 73.9%; Pred. No. 5.5e-121;

Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVGINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTV 87



Db 1 M V V K V G I N G F G R I G R L A F R R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60

QY 88 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F D V N G K F I K V S A E 147

Db 61 K D G G F E V N G N F I K V S A E X D P E ----- 81

QY 148 K D P E Q I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E 207

Db 82 -----N I D W A T D G V E I V L E 95

QY 208 A T S F F A K K E A A E K H L H A N G A K K V V I T A P G G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 267

Db 96 A T G F F A K K A A A E K H L H A N G A K K V V I T A P G G D D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 155

QY 268 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 327

Db 156 C L A P M A K A L Q D N F G V K Q G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A S N I V P N S T G A A 215

QY 328 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 387

Db 216 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V A V L E K E T S V E E I N A A M K A A A N D S Y G Y T E 275

QY 388 D P I V S S D I V G S V G S L F D A T Q T K M E V D G S Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 447

Db 276 D P I V S S D I I G M A Y G S L F D A T Q T K V Q T V D G N Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 335

QY 448 K 448

Db 336 K 336

RESULT 11

AAM50641

ID AAM50641 standard; protein; 336 AA.

AC AAM50641;

DT 04-APR-2002 (first entry)

DE Streptococcus uberis gapC plasmin binding protein.

XX GapC; plasmin-binding protein; UberGapC; infection; mastitis; vaccine;

XX diagnosis; therapy.

OS Streptococcus uberis.

PN WO200196381-A2.

PD 20-DEC-2001.

PF 11-JUN-2001; 2001WO-CA000838.

PR 12-JUN-2000; 2000US-0211022P.

XX (UYSA-) UNIV SASKATCHEWAN.

PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;

DR WPI; 2002-130725/17.

DR N-PSDB; ABA91250.

XX Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,

PT S.uberis, S.parauberis, or S.iniae, useful as vaccine component for

PT treating streptococcal infection which causes mastitis in vertebrates.

PS Claim 1(c); Fig 3A-B; 107pp; English.

XX The present sequence is that of the GapC plasmin binding protein

CC (UberGapC) of Streptococcus uberis ATCC 9927, encoded by the GapC gene

CC given in ABA91250. GapC, which has no signal sequence or membrane anchor

CC domain, is capable of eliciting an immune response in a vertebrate. The

CC invention provides the GapC genes and proteins of 5 Streptococcus

CC species, as well as recombinant vectors, host cells and vaccine

CC compositions comprising GapC polynucleotides or proteins. The vaccines

CC are used to treat or prevent a bacterial infection, especially a

CC streptococcal infection, and mastitis in particular (claimed). GapC

CC proteins are also used in claimed methods for detecting GapC antibodies,

CC and to raise antibodies that are used in claimed methods for detecting

CC GapC proteins. S. uberis is a common pathogen associated with mastitis in

CC cattle, horse, sheep and goat

XX

SQ Sequence 336 AA;

Query Match 68.5%; Score 1559.5; DB 5; Length 336;

Best Local Similarity 73.9%; Pred. NO. 5.5e-121;

Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 M V V K V G I N G F G R I G R L A F R R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 87

Db 1 M V V K V G I N G F G R I G R L A F R R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60

QY 88 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F D V N G K F I K V S A E 147

Db 61 K D G G F E V N G N F I K V S A E K D P E ----- 81

QY 148 K D P E Q I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E 207

Db 82 -----N I D W A T D G V E I V L E 95

QY 208 A T S F F A K K E A A E K H L H A N G A K K V V I T A P G G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 267

Db 96 A T G F F A K K A A A E K H L H A N G A K K V V I T A P G G D D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 155

QY 268 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 327

Db 156 C L A P M A K A L Q D N F G V K Q G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A S N I V P N S T G A A 215

QY 328 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 387

Db 216 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V A V L E K E T S V E E I N A A M K A A A N D S Y G Y T E 275

QY 388 D P I V S S D I V G S V G S L F D A T Q T K M E V D G S Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 447

Db 276 D P I V S S D I I G M A Y G S L F D A T Q T K V Q T V D G N Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 335

QY 448 K 448

Db 336 K 336

RESULT 12

AAM50666

ID AAM50666 standard; protein; 336 AA.

AC AAM50666;

DT 08-APR-2002 (first entry)

DE Streptococcus agalactiae gapC plasmin binding protein AgalGapC.

XX AgalGapC; GapC; plasmin binding protein; epitope; infection; vaccine;

XX immunisation; mastitis; therapy.

OS Streptococcus agalactiae.

PN WO200196379-A2.

XX 20-DEC-2001.

PF 11-JUN-2001; 2001WO-CA000836.

PR 12-JUN-2000; 2000US-0211247P.

XX (UYSA-) UNIV SASKATCHEWAN.

PI Potter AA, Perez-Casal J, Fontaine M;

XX WPI; 2002-098051/13.  
DR N-PSDB; ABA91329.  
XX Novel GapC multiple epitope fusion polypeptide comprising antigenic  
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.  
PT parauberis, or S. iniae GapC protein useful for treating mastitis in  
PT vertebrates.  
XX  
PS Claim 8; Fig 2A-B; 116pp; English.  
XX  
CC The present sequence is that of the GapC plasmin binding protein,  
CC AgalGapC, of Streptococcus agalactiae ATCC 27541. The invention relates  
CC to novel GapC multiple epitope fusion proteins that comprise epitopes  
CC from 1 or more of Streptococcus dysgalactiae, Streptococcus agalactiae,  
CC Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae  
CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).  
CC Expression vectors and host cells for production of the multiple epitope  
CC fusion proteins are provided. The multiple epitope proteins are used in  
CC claimed vaccines for treating or preventing a bacterial infection in a  
CC vertebrate, especially a streptococcal infection, and particularly  
CC mastitis. They are also used in claimed methods of detecting  
CC Streptococcus antibodies. The multiple epitope proteins are capable of  
CC eliciting broad immunity against a variety of streptococcal infections  
CC while minimising the number of antigens present in the final formulation  
CC and concomitantly reducing production costs  
XX  
SQ Sequence 336 AA;  
  
Query Match 68.4%; Score 1557.5; DB 5; Length 336;  
Best Local Similarity 73.9%; Pred. No. 8.1e-121;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;  
  
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRF----- 54  
  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEIVLEALEGTVEVKDGGFDVNGKFIKVS 147  
Db 55 ----- 54  
  
QY 148 KDPEQIDWATDGVLEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIV 207  
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 95  
  
QY 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 155  
  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAIVPNSTGAA 215  
  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKNVSVDEINAAKAAANDSFYGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLLEKDVTVVEVNAAMKAAANDSFYGYTE 275  
  
QY 388 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLLEYFAKIA 447  
Db 276 DPIVSSDIVGSYGLFDATQTKVQTVDNQNLVKVSVWYDNEMSYTSQLVRLLEYFAKIA 335  
  
QY 448 K 448  
Db 336 K 336  
  
RESULT 13  
AAM50640  
ID AAM50640 standard; protein; 336 AA.  
XX  
AC AAM50640;  
XX  
DT 04-APR-2002 (first entry)

XX Streptococcus agalactiae gapC plasmin binding protein.  
DE  
XX GapC; plasmin-binding protein; AgalGapC; infection; mastitis; vaccine;  
KW diagnosis; therapy.  
KW  
XX Streptococcus agalactiae.  
OS  
XX WO200196381-A2.  
PN  
XX 20-DEC-2001.  
PD  
XX 11-JUN-2001; 2001WO-CA000838.  
PF  
XX 12-JUN-2000; 2000US-0211022P.  
PR  
XX (UYSA-) UNIV SASKATCHEWAN.  
PA  
PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;  
XX WPI; 2002-130725/17.  
DR N-PSDB; ABA91249.  
DR  
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,  
PT S. uberis, S. parauberis, or S. iniae, useful as vaccine component for  
PT treating streptococcal infection which causes mastitis in vertebrates.  
PT  
XX Claim 1(b); Fig 2A-B; 107pp; English.  
PS  
XX The present sequence is that of the GapC plasmin binding protein  
CC (AgalGapC) of Streptococcus agalactiae ATCC 27541, encoded by the GapC  
CC gene given in ABA91249. GapC, which has no signal sequence or membrane  
CC anchor domain, is capable of eliciting an immune response in a  
CC vertebrate. The invention provides the GapC genes and proteins of 5  
CC Streptococcus species, as well as recombinant vectors, host cells and  
CC vaccine compositions comprising GapC polynucleotides or proteins. The  
CC vaccines are used to treat or prevent a bacterial infection, especially a  
CC streptococcal infection, and mastitis in particular (claimed). GapC  
CC proteins are also used in claimed methods for detecting GapC antibodies,  
CC and to raise antibodies that are used in claimed methods for detecting  
CC GapC proteins. S. agalactiae is a common pathogen associated with  
CC mastitis in cattle, horse, sheep and goat. It also causes septicemia,  
CC meningitis, bacteraemia, impetigo, arthritis, urinary tract infections,  
CC abscesses, spontaneous abortion, etc  
XX  
SQ Sequence 336 AA;  
  
Query Match 68.4%; Score 1557.5; DB 5; Length 336;  
Best Local Similarity 73.9%; Pred. No. 8.1e-121;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;  
  
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRF----- 54  
  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEIVLEALEGTVEVKDGGFDVNGKFIKVS 147  
Db 55 ----- 54  
  
QY 148 KDPEQIDWATDGVLEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIV 207  
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 95  
  
QY 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 155  
  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAIVPNSTGAA 215  
  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKNVSVDEINAAKAAANDSFYGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLLEKDVTVVEVNAAMKAAANDSFYGYTE 275

Db 216 KAIGLVIPELNGKLDGAAQRVPVPTGTSVTELVALEKDVTEEVNAAMKAANDSYGYTE 275

QY 388 DPIVSSDIVGVSYSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 276 DPIVSSDIVGISYGLFDATQTKVQTVDGNQLVKVSWYDNEMSYTSQLVRTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 14

ABU02516

ID ABU02516 standard; protein; 335 AA.

XX

AC ABU02516;

XX

DT 23-OCT-2003 (revised)

DT 11-FEB-2003 (first entry)

XX

DE S. pneumoniae type 4 strain protein from coding region #2094.

DE

KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;

KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;

KW gene therapy; vaccine.

XX

OS Streptococcus pneumoniae; type 4 strain.

XX

PN WO200277021-A2.

XX

PD 03-OCT-2002.

XX

XX 27-MAR-2002; 2002WO-IB002163.

PF

XX 27-MAR-2001; 2001GB-00007658.

PR

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX

PI Masignani V, Tettelin H, Fraser C;

XX

DR WPI; 2003-040579/03.

DR N-PSDB; ABX07806.

XX

PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,

PT useful as medicaments for treating or preventing a disease or infection

PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or

PT ear infection.

XX

PS Claim 1; SEQ ID NO 4188; 56pp; English.

XX

CC The invention relates to a protein comprising or having at least 50%

CC identity to any of the 2469 amino acid sequences, identified in the

CC specification (available on a computer readable format), or its fragment,

CC expressed from 2469 of 2489 identified DNA coding regions from the

CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as

CC ABS56454. Also included are an antibody which binds one of the proteins,

CC treating a patient by administering the protein, DNA or antibody (in a

CC composition), a kit comprising first and second primers, which are the

CC nucleic acid cited above or fragments between nucleotides 8-100 of a

CC sequence not defined in the specification, for amplifying a target

CC sequence contained within a Streptococcus nucleic acid sequence, where

CC the first primer is substantially complementary to the target sequence

CC and the second primer is substantially complementary to the complement of

CC the target sequence, and where the parts of the primers having

CC substantial complementarity define the termini of the target sequence to

CC be amplified, assay comprising contacting a test compound with the

CC protein, and determining whether the test compound binds to the protein

CC and a Streptococcus pneumoniae bacterium, where one or more genes

CC encoding the proteins has been rendered inactive. The proteins, nucleic

CC acid molecules, antibody and compositions are useful as medicaments for

CC treating or preventing a disease or infection due to streptococcus

CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis

CC media or ear infection. They are also useful in developing vaccines,

CC diagnostics and antibiotics. The methods are useful for identifying

CC immunodominant proteins. The present sequence is one of the 2469 proteins

CC expressed by the identified coding regions from the genomic sequence.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 23-OCT-2003 to

CC standardise OS field)

XX

SQ Sequence 335 AA;

Query Match 67.4%; Score 1535; DB 6; Length 335;

Best Local Similarity 73.6%; Pred. No. 6e-119;

Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVETRINDLTDPNMLAHLKDYDTQGRFDGTVEV 87

Db 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVETRINDLTDPNMLAHLKDYDTQGRFDGTVEV 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147

Db 61 KEGGFEVNGKFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207

Db 82 -----QIDWATDGVVEIVLE 95

QY 208 ATFFFAKKEAAEKHLHANGAKKVITAPGNDVKTVVENTNHDILDTETVISGASCTTN 267

Db 96 ATGFFFAKKEAAEKHL-KGGAKKVITAPGNDVKTVVENTNHDVLDGTETVISGASCTTN 154

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMI LDGPHRGDLRRARAGAAANIVPNSTGAA 327

Db 155 CLAPMAKALQDNFVGVEGLMTTIHAYTGDQMI LDGPHRGDLRRARAGAAANIVPNSTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQRVPVPTGTSVTELVTLDKNVSVDEINAAMKAANDSFGYTE 387

Db 215 KAIGLVIPELNGKLDGSAQRVPTGTSVTELVAVLEKNVTVDENAMKAASNESYGYTE 274

QY 388 DPIVSSDIVGVSYSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 275 DPIVSSDIVGMSYGLFDATQTKVLDVDGKQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 334

QY 448 K 448

Db 335 K 335

RESULT 15

AAU37576

ID AAU37576 standard; protein; 359 AA.

XX

AC AAU37576;

DT 14-FEB-2002 (first entry)

XX

DE Streptococcus pneumoniae cellular proliferation protein #5.

XX

KW Antisense; prokaryotic cellular proliferation protein; antibiotic;

KW antibacterial; drug design.

XX

OS Streptococcus pneumoniae.

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-US009180.

XX

PR 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.



PR 27-NOV-2000; 2000US-0253625P.  
PR 22-DEC-2000; 2000US-0257931P.  
PR 16-FEB-2001; 2001US-0269308P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
XX  
DR WPI; 2001-611495/70.  
DR N-PSDB; AAS55435.  
XX  
PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.  
XX  
PS Example 3; SEQ ID NO 13169; 511pp; English.  
XX

Search completed: September 15, 2004, 10:24:29  
Job time : 131 secs

CC The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes  
CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence represents an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 359 AA;

Query Match 67.4%; Score 1535; DB 4; Length 359;  
Best Local Similarity 73.6%; Pred. No. 6.6e-119;  
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVGINGFGRGLAFRRIONVEGVETRLNDLTDPMHLALLKYDTTQGRFDGTVEV 87  
Db |||||||  
25 MVVKVGINGFGRGLAFRRIONVEGVETRLNDLTDPMHLALLKYDTTQGRFDGTVEV 84  
QY 88 KEGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
Db |||||||  
85 KEGFEVNGKFIKVSARDPE----- 105  
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 207  
Db |||||||  
106 -----QIDWATDGVLEIVLE 119  
QY 208 ATSFFAKKEAAEKHLHANGAKKWITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db |||||||  
120 ATGFFAKKEAAEKHL-KGAKKWITAPGNDVKTVVFNTHDVLDTGTETVISGASCTTN 178  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPFRGGDLRRARAGAAIVPNSTGAA 327  
Db |||||||  
179 CLAPMAKALQDNFVVEGLMTTIHAYTGDQMLDGPFRGGDLRRARAGAAIVPNSTGAA 238  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLDKNVSVDEINAAMKAASNDSPGYTE 387  
Db |||||||  
239 KAIGLVIPELNGKLDGSAQRPVPTGTSVTELVAVLEKNVTVDENVNAAMKAASNESYGYTE 298  
QY 388 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNMSYTAQLVRTLEYFAKIA 447  
Db |||||||  
299 DPIVSSDIVGSYGLFDATQTKVLDVGKQLVKVSVWYDNMSYTAQLVRTLEYFAKIA 358  
QY 448 K 448  
Db |  
359 K 359

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:00:07 ; Search time 34 Seconds  
(without alignments)  
680.248 Million cell updates/sec

Title: US-10-650-369-22  
Perfect score: 2278  
Sequence: 1 MKKITGIILLALLAVIILSAC.....EMSYTAQLVRLTLEYFAKIAK 448

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2278	100.0	448	US-09-878-766A-22	Sequence 22, Appl
2	1656.5	72.7	336	US-09-878-766A-12	Sequence 12, Appl
3	1652.5	72.5	336	US-07-928-462-2	Sequence 2, Appli
4	1652.5	72.5	336	US-08-273-247-2	Sequence 2, Appli
5	1559.5	68.5	336	US-09-878-766A-16	Sequence 16, Appl
6	1557.5	68.4	336	US-09-878-766A-14	Sequence 14, Appl
7	1534.5	67.4	336	US-09-878-766A-20	Sequence 20, Appl
8	1530.5	67.2	336	US-09-878-766A-18	Sequence 18, Appl
9	1491	65.5	333	US-08-961-083-54	Sequence 54, Appl
10	1491	65.5	333	US-09-536-784-54	Sequence 54, Appl
11	1282	56.3	346	US-09-134-000C-4400	Sequence 4400, Ap
12	1270	55.8	333	US-09-107-532A-4769	Sequence 4769, Ap
13	1138.5	50.0	348	US-09-134-001C-5513	Sequence 5513, Ap
14	923.5	40.5	357	US-09-134-000C-4229	Sequence 4229, Ap
15	918.5	40.3	340	US-09-634-238-233	Sequence 233, App
16	905.5	39.7	349	US-09-489-039A-13939	Sequence 13939, A
17	849	37.3	335	5290690-9	Patent No. 5290690
18	838	36.8	334	5290690-10	Patent No. 5290690
19	803	35.3	340	US-09-095-855-205	Sequence 205, App
20	803	35.3	340	US-09-205-426-205	Sequence 205, App
21	780	34.2	334	5290690-11	Patent No. 5290690
22	776.5	34.1	335	US-09-489-039A-7679	Sequence 7679, Ap
23	758.5	33.3	282	US-09-107-532A-4057	Sequence 4057, Ap
24	747.5	32.8	340	US-09-543-681A-8083	Sequence 8083, Ap
25	719	31.6	352	US-09-489-039A-9410	Sequence 9410, Ap
26	690	30.3	336	US-09-198-452A-662	Sequence 662, App
27	683.5	30.0	338	US-09-091-725-10	Sequence 10, Appl

28	678	29.8	268	2	US-08-997-080-186	Sequence 186, App
29	678	29.8	268	2	US-08-997-362-186	Sequence 186, App
30	678	29.8	268	3	US-09-095-855-186	Sequence 186, App
31	678	29.8	268	4	US-09-324-542-186	Sequence 186, App
32	678	29.8	268	4	US-09-205-426-186	Sequence 186, App
33	675	29.6	335	2	US-08-903-800A-6	Sequence 6, Appli
34	668.5	29.3	340	4	US-09-252-991A-21612	Sequence 21612, A
35	660	29.0	334	4	US-09-674-826B-6	Sequence 6, Appli
36	646	28.4	341	4	US-09-134-001C-3783	Sequence 3783, Ap
37	628	27.6	333	3	US-09-532-803-11	Sequence 11, Appl
38	617	27.1	333	3	US-09-532-803-9	Sequence 9, Appli
39	601.5	26.4	277	4	US-09-548-938A-13	Sequence 13, Appl
40	542.5	23.8	442	4	US-09-540-236-3378	Sequence 3378, Ap
41	540.5	23.7	350	4	US-09-489-039A-11854	Sequence 11854, A
42	535	23.5	340	4	US-09-328-352-6704	Sequence 6704, Ap
43	532.5	23.4	417	4	US-09-540-236-3470	Sequence 3470, Ap
44	512.5	22.5	359	4	US-09-543-681A-6082	Sequence 6082, Ap
45	504.5	22.1	490	4	US-09-328-352-4616	Sequence 4616, Ap

ALIGNMENTS

RESULT 1  
US-09-878-766A-22  
; Sequence 22, Application US/09878766A  
; Patent No. 6660270  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 448  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple  
; OTHER INFORMATION: epitope fusion protein  
US-09-878-766A-22

Query Match		100.0%;	Score 2278;	DB 4;	Length 448;
Best Local Similarity		100.0%;	Pred. No. 1.7e-206;		
Matches 448;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKKITGIILLALLAVIILSACQANYGSGMVVKVINGFGRIGRLAFRRIONVEGVEVTRIN	60		
Db	1	MKKITGIILLALLAVIILSACQANYGSGMVVKVINGFGRIGRLAFRRIONVEGVEVTRIN	60		
QY	61	DLTDPNMLAHLKYDITTOGRFDGTVEVKEGGFEVNGNFIKVSAERDPENIDWATDGV	120		
Db	61	DLTDPNMLAHLKYDITTOGRFDGTVEVKEGGFEVNGNFIKVSAERDPENIDWATDGV	120		
QY	121	LEALEGTVEVKDGGFDVNGKFIKVSAREKDEQIDWATDGV	180		
Db	121	LEALEGTVEVKDGGFDVNGKFIKVSAREKDEQIDWATDGV	180		
QY	181	GQFVKVSAEREPAIDWATDGV	240		
Db	181	GQFVKVSAEREPAIDWATDGV	240		
QY	241	KTVVNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTHAYTGDMIL	300		
Db	241	KTVVNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTHAYTGDMIL	300		
QY	301	DGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTGCVTEL	360		



Db 301 DPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTGVTTELVV 360  
QY 361 TLDKNVSVDEINAAKKAASNDSEFGYTEDPIVSSDIVGVSYGSLFDATQTKVMEVDGSQLV 420  
Db 361 TLDKNVSVDEINAAKKAASNDSEFGYTEDPIVSSDIVGVSYGSLFDATQTKVMEVDGSQLV 420  
QY 421 KVVSWYDNEMSYTAQLVRTLEYFAKIAK 448  
Db 421 KVVSWYDNEMSYTAQLVRTLEYFAKIAK 448

RESULT 2  
US-09-878-766A-12  
; Sequence 12, Application US/09878766A  
; Patent No. 6660270  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus dysgalactiae  
US-09-878-766A-12

Query Match 72.7%; Score 1656.5; DB 4; Length 336;  
Best Local Similarity 79.6%; Pred. No. 5e-148;  
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;  
QY 28 MVVKVGINGFGRIGRLAFRRIQNVGEVETRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 87  
Db 1 MVVKVGINGFGRIGRLAFRRIQNVGEVETRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVVEKDGFDVNGKFIKVSAAE 147  
Db 61 KEGGFEVNGNFIKVSARDPE-----81

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207  
Db -----NIDWATDGVLE 95  
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVPVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVPVFNTHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAIVPNSTGAA 327  
Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLTKVSVDEINAAKKAASNDSEFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLTKVSVDEINAAKKAASNDSEFGYTE 275  
QY 388 DPIVSSDIVGVSYGSLFDATQTKVMEVDGSQLKVVSWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGVSYGSLFDATQTKVMEVDGSQLKVVSWYDNEMSYTAQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336

RESULT 3  
US-07-928-462-2  
; Sequence 2, Application US/07928462  
; Patent No. 5328996

; GENERAL INFORMATION:  
; APPLICANT: Boyle, Michael D.P.  
; APPLICANT: Lottenberg, Richard  
; APPLICANT: Broder, Christopher C.  
; APPLICANT: von Mering, Gregory O.  
; TITLE OF INVENTION: Bacterial Plasmin Receptors as  
; TITLE OF INVENTION: Fibrinolytic Agents  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/928,462  
; FILING DATE: 19920810  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/524,411  
; FILING DATE: 16-MAY-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/330,849  
; FILING DATE: 29-MAR-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-928-462-2

Query Match 72.5%; Score 1652.5; DB 1; Length 336;  
Best Local Similarity 79.1%; Pred. No. 1.2e-147;  
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;  
QY 28 MVVKVGINGFGRIGRLAFRRIQNVGEVETRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 87  
Db 1 MVVKVGINGFGRIGRLAFRRIQNVGEVETRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVVEKDGFDVNGKFIKVSAAE 147  
Db 61 KEGGFEVNGNFIKVSARDPE-----81  
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207  
Db -----NIDWATDGVLE 95  
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVPVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVPVFNTHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAIVPNSTGAA 327  
Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLTKVSVDEINAAKKAASNDSEFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLTKVSVDEINAAKKAASNDSEFGYTE 275

QY 388 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336  
RESULT 4  
US-08-273-247-2  
; Sequence 2, Application US/08273247  
; Patent No. 6136323  
; GENERAL INFORMATION:  
; APPLICANT: Boyle, Michael D.P.  
; APPLICANT: Lottenberg, Richard  
; APPLICANT: Broder, Christopher C.  
; APPLICANT: von Mering, Gregory O.  
; TITLE OF INVENTION: Bacterial Plasmin Receptors as  
; TITLE OF INVENTION: Fibrinolytic Agents  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/273,247  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/524,411  
; FILING DATE: 16-MAY-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/330,849  
; FILING DATE: 29-MAR-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-273-247-2  
Query Match 72.5%; Score 1652.5; DB 3; Length 336;  
Best Local Similarity 79.1%; Pred. No. 1.2e-147;  
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;  
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKDYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONIEGVETRLNDLTPNMLAHLKDYDTTQGRFDGTVEV 60  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVS 147  
Db 61 KEGGFEVNGNFIKVSARDPE-----81  
QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207  
Db 82 -----NIDWATDGVVEIVLE 95

QY 208 ATSPFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAANIVPNSTGAA 327  
Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAANIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVTLDKNVSVDEINAAKAAASNDSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVTLDKNVSVDEINAAKAAASNDSFGYTE 275  
QY 388 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336  
RESULT 5  
US-09-878-766A-16  
; Sequence 16, Application US/09878766A  
; Patent No. 6660270  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus uberis  
; US-09-878-766A-16  
Query Match 68.5%; Score 1559.5; DB 4; Length 336;  
Best Local Similarity 73.9%; Pred. No. 7.1e-139;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;  
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKDYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKDYDTTQGRFDGTVEV 60  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVS 147  
Db 61 KEGGFEVNGNFIKVSARDPE-----81  
QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207  
Db 82 -----NIDWATDGVVEIVLE 95  
QY 208 ATSPFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGDDVKTVVFNTHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAANIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGASNIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVTLDKNVSVDEINAAKAAASNDSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVALEKETSVEEINAAKAAANDSYGYTE 275  
QY 388 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 335

Db 276 DPIVSSDIIGMAYGSLFDATQTKVQTVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336

RESULT 6

US-09-878-766A-14  
; Sequence 14, Application US/09878766A  
; Patent No. 6660270  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus agalactiae  
US-09-878-766A-14

Query Match 68.4%; Score 1557.5; DB 4; Length 336;  
Best Local Similarity 73.9%; Pred. No. 1.1e-138;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 87  
Db 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPNMLAHLKLYDTTQGRF----- 54  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVS AE 147  
Db 55 ----- 54  
QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207  
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 95  
QY 208 ATSFFAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFASKAKAGQHIHENGAKKVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAPGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAANIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMILDGPHRGDLRRARAGAANIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVTTELVTLDKNVSVDEINAAKKAASNDSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVTTELVTLEKDVTEEVEVNAAMKAAANDSYGYTE 275  
QY 388 DPIVSSDIVGSYGS LFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGISYGS LFDATQTKVQTVDGNQLVKVSWYDNEMSYTSQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336

RESULT 7

US-09-878-766A-20  
; Sequence 20, Application US/09878766A  
; Patent No. 6660270  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus iniae  
US-09-878-766A-20

Query Match 67.4%; Score 1534.5; DB 4; Length 336;  
Best Local Similarity 72.7%; Pred. No. 1.6e-136;  
Matches 306; Conservative 17; Mismatches 13; Indels 85; Gaps 1;  
QY 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 87  
Db 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVS AE 147  
Db 61 KDGGEVNGS----- 70

QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207  
Db 71 -----FVKVSAEREPANIDWATDGV EIVLE 95  
QY 208 ATSFFAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFASKAAAEQHIHANGAKKVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAPGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAANIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHGYTGDQMVLDPHRRGDLRRARAAAANIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVTTELVTLDKNVSVDEINAAKKAASNDSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVTTELVAVLEKDTSV EEAAMKAAANDSYGYTE 275  
QY 388 DPIVSSDIVGSYGS LFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DAIVSSDIVGISYGS LFDATQTKVQTVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336

RESULT 8

US-09-878-766A-18  
; Sequence 18, Application US/09878766A  
; Patent No. 6660270  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus parauberis  
US-09-878-766A-18

Query Match 67.2%; Score 1530.5; DB 4; Length 336;  
Best Local Similarity 73.4%; Pred. No. 3.9e-136;



Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;  
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFD----- 55  
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
Db 56 -----GTVEVKDGGFDVNGKFIKVSAAE 77  
QY 148 KOPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207  
Db 78 KDPEQ-----IDWATDGVIEVLE 95  
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFAKKAAAEKHLHENGAKKVVITAPGGDDVKTVVFNTHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDDGPHRGDLRRARAGAAANIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMILDDGPHRGDLRRARAGANNIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSEGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVAVLNKETSVEEINSMKAAANDSEGYTE 275  
QY 388 DPIVSSDIVGVSGLFDATQTKMVEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGMSFGLFDATQTKVQTVDGNQLVKVSVWYDNEMSYTAQLDRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336

RESULT 9  
US-08-961-083-54  
; Sequence 54, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,083  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 333 amino acids  
; TYPE: amino acid

STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-961-083-54  
Query Match 65.5%; Score 1491; DB 3; Length 333;  
Best Local Similarity 72.4%; Pred. No. 2.1e-132;  
Matches 302; Conservative 12; Mismatches 17; Indels 86; Gaps 2;  
QY 29 VVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEVK 88  
Db 1 VVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTVEVK 60  
QY 89 EGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAAE 148  
Db 61 EGGFEVNGKFIKVSARDPE----- 80  
QY 149 DPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLEA 208  
Db 81 -----QIDWATDGVIEVLEA 95  
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTNC 268  
Db 96 TGFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVVFNTHDVLDTETVISGASCTTNC 154  
QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDDGPHRGDLRRARAGAAANIVPNSTGAAK 328  
Db 155 LAPMAKALQDNFVGVEGLMTTIHAYTGDQMILDDGPHRGDLRRARAGAAANIVPNSTGAAK 214  
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSEGYTED 388  
Db 215 AIGLVIPELNGKLDGSAQRPVPTGVSVELVAVLEKNVTVDENVAAKKAASNEGYTED 274  
QY 389 PIVSSDIVGVSGLFDATQTKMVEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAK 445  
Db 275 PIVSSDIVGMSYGLFDATQTKVLDVKGQLVKVSVWYDNEMSYTAQLVRTLGILRK 331

RESULT 10  
US-09-536-784-54  
; Sequence 54, Application US/09536784  
; Patent No. 6573082  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/536,784  
; FILING DATE: 30-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: OCT-30-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41,971  
; REFERENCE/DOCKET NUMBER: PB340P3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 54:





QY 445 KI 446  
Db 356 NL 357

RESULT 15

US-09-634-238-233  
; Sequence 233, Application US/09634238  
; Patent No. 6544772

GENERAL INFORMATION:

; APPLICANT: Glenn, Matthew  
; APPLICANT: Havukkala, Ilkka J.  
; APPLICANT: Bloksberg, Leonard, N.  
; APPLICANT: Lubbers, Mark W.  
; APPLICANT: Dekker, James  
; APPLICANT: Christensson, Anna C.  
; APPLICANT: Holland, Ross  
; APPLICANT: O'Toole, Paul W.  
; APPLICANT: Reid, Julian R.  
; APPLICANT: Coolbear, Timothy  
; TITLE OF INVENTION: Polynucleotides, materials incorporating  
; TITLE OF INVENTION: them and methods for using them.

; FILE REFERENCE: 11000.1043U1  
; CURRENT APPLICATION NUMBER: US/09/634,238

; CURRENT FILING DATE: 2000-08-08

; NUMBER OF SEQ ID NOS: 422

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 233

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Lactobacillus rhamnosus

US-09-634-238-233

Query Match 40.3%; Score 918.5; DB 4; Length 340;  
Best Local Similarity 46.0%; Pred. NO. 2.3e-78;  
Matches 196; Conservative 49; Mismatches 88; Indels 93; Gaps 7;

QY 28 M V K V G I N G F G R I G R L A F R R I Q N V - - - - E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G 83  
Db 1 M T V K I G I N G F G R I G R L A F R R I Y E I G A K S N D I Q V V A I N D L T S P T M L A H L L K Y D S T H G T F P G 60  
QY 84 T V E V K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F D V N G K F I K 143  
Db 61 E V S A T D N G I V V D G K E Y R V Y A E P Q A Q N I P W - - - - - V K N - - - - - 92  
QY 144 V S A E K D P E Q I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E 203  
Db 93 - - - - - 96  
QY 204 I V L E A T S F F A K K E A E K H L H A N G A K K V V I T A P G G N D V K T V V E N T N H D I L D G T E T V I S G A S 263  
Db 97 Y V L E C T G F T S A E K S Q A H L D A - G A K R V L I S A P A G K - I K T I V Y N V N D D T L N A D D K I V S A G S 154  
QY 264 C T T N C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A G A A N I V P N S 323  
Db 155 C T T N C L A P M A Y F L N Q E F G I E V G T M T V H A Y T S T Q M L L D G P V R G G N L R A A R S A A A N T I P H S 214  
QY 324 T G A A K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V V T L - D K N V S V D E I N A A M K - - A A S N 380  
Db 215 T G A A K A I G L V I P E L N G K L Q H A Q R V S V V D G S L T E L V S I L K T K N V T A D Q V N E A I K K H T E N N 274  
QY 381 D S F G Y T E D P I V S S D I V G S Y G S L F D A T Q T K V M E V D G S Q L V K V S W Y D N E M S Y T A Q L V R T L 440  
Db 275 P S F G N E D E I V S S D V I G T T Y G S I F D T Q T E V T T A G D Y Q L V K T V A W Y D N E Y G F T C Q M I R T L 334  
QY 441 E Y F A K I 446  
Db 335 L K F A T L 340



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:08:58 ; Search time 131 Seconds  
(without alignments)  
1096.710 Million cell updates/sec

Title: US-10-650-369-22  
Perfect score: 2278  
Sequence: 1 MKKITGIILLLLAVIILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %		Length	DB	ID	Description
	Score					
1	2278	100.0	448	9	US-09-878-766A-22	Sequence 22, Appl
2	2278	100.0	448	12	US-10-650-369-22	Sequence 22, Appl
3	1656.5	72.7	336	9	US-09-878-766A-12	Sequence 12, Appl
4	1656.5	72.7	336	10	US-09-878-781-4	Sequence 4, Appli
5	1656.5	72.7	336	12	US-10-650-369-12	Sequence 12, Appl
6	1656.5	72.7	336	14	US-10-134-297-4	Sequence 4, Appli
7	1655.5	72.7	336	12	US-10-282-122A-74379	Sequence 74379, A
8	1650	72.4	335	10	US-09-878-781-14	Sequence 14, Appl
9	1562.5	68.6	336	10	US-09-878-781-16	Sequence 16, Appl
10	1559.5	68.5	336	9	US-09-878-766A-16	Sequence 16, Appl
11	1559.5	68.5	336	10	US-09-878-781-8	Sequence 8, Appli
12	1559.5	68.5	336	12	US-10-650-369-16	Sequence 16, Appl
13	1559.5	68.5	336	14	US-10-134-297-8	Sequence 8, Appli
14	1557.5	68.4	336	9	US-09-878-766A-14	Sequence 14, Appl
15	1557.5	68.4	336	10	US-09-878-781-6	Sequence 6, Appli

16	1557.5	68.4	336	12	US-10-650-369-14	Sequence 14, Appl
17	1557.5	68.4	336	14	US-10-134-297-6	Sequence 6, Appli
18	1535	67.4	359	9	US-09-815-242-13169	Sequence 13169, A
19	1535	67.4	359	12	US-10-282-122A-74186	Sequence 74186, A
20	1534.5	67.4	336	9	US-09-878-766A-20	Sequence 20, Appl
21	1534.5	67.4	336	10	US-09-878-781-12	Sequence 12, Appl
22	1534.5	67.4	336	12	US-10-650-369-20	Sequence 20, Appl
23	1534.5	67.4	336	14	US-10-134-297-12	Sequence 12, Appl
24	1534	67.3	359	9	US-09-815-242-13593	Sequence 13593, A
25	1530.5	67.2	336	9	US-09-878-766A-18	Sequence 18, Appl
26	1530.5	67.2	336	10	US-09-878-781-10	Sequence 10, Appl
27	1530.5	67.2	336	12	US-10-650-369-18	Sequence 18, Appl
28	1530.5	67.2	336	14	US-10-134-297-10	Sequence 10, Appl
29	1521	66.8	337	12	US-10-282-122A-72254	Sequence 72254, A
30	1491	65.5	333	9	US-09-765-272-54	Sequence 54, Appl
31	1357.5	59.6	336	15	US-10-369-493-18524	Sequence 18524, A
32	1282	56.3	333	9	US-09-815-242-10847	Sequence 10847, A
33	1282	56.3	333	12	US-10-282-122A-57294	Sequence 57294, A
34	1280	56.2	337	15	US-10-369-493-18342	Sequence 18342, A
35	1224.5	53.8	335	12	US-10-282-122A-53266	Sequence 53266, A
36	1205.5	52.9	336	12	US-10-282-122A-60492	Sequence 60492, A
37	1203	52.8	334	12	US-10-282-122A-51519	Sequence 51519, A
38	1166	51.2	335	12	US-10-282-122A-71734	Sequence 71734, A
39	1145	50.3	334	12	US-10-282-122A-65661	Sequence 65661, A
40	1141	50.1	357	12	US-10-282-122A-65001	Sequence 65001, A
41	1138.5	50.0	336	12	US-10-282-122A-70890	Sequence 70890, A
42	1132.5	49.7	336	9	US-09-815-242-5618	Sequence 5618, Ap
43	1132.5	49.7	336	9	US-09-815-242-12540	Sequence 12540, A
44	1132.5	49.7	336	12	US-10-282-122A-43997	Sequence 43997, A
45	941.5	41.3	333	12	US-10-282-122A-56828	Sequence 56828, A

ALIGNMENTS

RESULT 1  
US-09-878-766A-22  
; Sequence 22, Application US/09878766A  
; Patent No. US20020044928A1  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 448  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple  
; OTHER INFORMATION: epitope fusion protein  
US-09-878-766A-22

Query Match 100.0%; Score 2278; DB 9; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.2e-198;  
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKKITGIILLLLAVIILSACQANYGSMVVKVGINGFRIGRLAFRIQNVGEVETRIN	60
Db	1	MKKITGIILLLLAVIILSACQANYGSMVVKVGINGFRIGRLAFRIQNVGEVETRIN	60
QY	61	DLTDPNMLAHLKLYDTTQGRFDGTVEVKEGGFEVNGNFIKVSARDENIDWATDGEIV	120
Db	61	DLTDPNMLAHLKLYDTTQGRFDGTVEVKEGGFEVNGNFIKVSARDENIDWATDGEIV	120
QY	121	LEALEGTVEVKDGGFDVNGKFIKVSAEKDPEQIDWATDGEIVLEIDGTVEVKEGGFEVN	180

Db 121 LEAEGTVEVKDGGFDVNGKFIKVSAAEKDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVN 180  
QY 181 GQFVKVSAEREPANIDWATDGV EIVLEATSPFAKKEAAEKHLHANGAKKVITAPGGNDV 240  
Db 181 GQFVKVSAEREPANIDWATDGV EIVLEATSPFAKKEAAEKHLHANGAKKVITAPGGNDV 240  
QY 241 KTVVFNTHDILDGTETVISGASCTTNC LAPMAKALHDAFGIQKGLMTTIHAYTGDQMIL 300  
Db 241 KTVVFNTHDILDGTETVISGASCTTNC LAPMAKALHDAFGIQKGLMTTIHAYTGDQMIL 300  
QY 301 DGPHRGGDLRRARAGAAANIVPNSTGA AKAIGLVIPELNGKLDGAAQRPVPTG SVTELVV 360  
Db 301 DGPHRGGDLRRARAGAAANIVPNSTGA AKAIGLVIPELNGKLDGAAQRPVPTG SVTELVV 360  
QY 361 TLDKNVSVDEINAAKKAASNDSFGYT EDPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLV 420  
Db 361 TLDKNVSVDEINAAKKAASNDSFGYT EDPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLV 420  
QY 421 KVVSWYDNEMSYTAQLVRTLEYFAKIAK 448  
Db 421 KVVSWYDNEMSYTAQLVRTLEYFAKIAK 448

RESULT 2  
US-10-650-369-22  
; Sequence 22, Application US/10650369  
; Publication No. US20040062774A1  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/10/650,369  
; CURRENT FILING DATE: 2003-08-27  
; PRIOR APPLICATION NUMBER: US/09/878,766A  
; PRIOR FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 448  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple  
; OTHER INFORMATION: epitope fusion protein  
US-10-650-369-22

Query Match 100.0%; Score 2278; DB 12; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.2e-198;  
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKITGIILLLLAVIILSACQANYGSGMVVKVINGFG RIGRLAFRRIONVEGVEVTRIN 60  
Db 1 MKKITGIILLLLAVIILSACQANYGSGMVVKVINGFG RIGRLAFRRIONVEGVEVTRIN 60  
QY 61 DLTDPNMLAHLKDYDTTQGRFDGTVEVKEGGFEVNGFNFIKVS AERDPENIDWATDGV EIV 120  
Db 61 DLTDPNMLAHLKDYDTTQGRFDGTVEVKEGGFEVNGFNFIKVS AERDPENIDWATDGV EIV 120  
QY 121 LEAEGTVEVKDGGFDVNGKFIKVSAAEKDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVN 180  
Db 121 LEAEGTVEVKDGGFDVNGKFIKVSAAEKDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVN 180  
QY 181 GQFVKVSAEREPANIDWATDGV EIVLEATSPFAKKEAAEKHLHANGAKKVITAPGGNDV 240  
Db 181 GQFVKVSAEREPANIDWATDGV EIVLEATSPFAKKEAAEKHLHANGAKKVITAPGGNDV 240  
QY 241 KTVVFNTHDILDGTETVISGASCTTNC LAPMAKALHDAFGIQKGLMTTIHAYTGDQMIL 300  
Db 241 KTVVFNTHDILDGTETVISGASCTTNC LAPMAKALHDAFGIQKGLMTTIHAYTGDQMIL 300

QY 301 DGPHRGGDLRRARAGAAANIVPNSTGA AKAIGLVIPELNGKLDGAAQRPVPTG SVTELVV 360  
Db 301 DGPHRGGDLRRARAGAAANIVPNSTGA AKAIGLVIPELNGKLDGAAQRPVPTG SVTELVV 360  
QY 361 TLDKNVSVDEINAAKKAASNDSFGYT EDPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLV 420  
Db 361 TLDKNVSVDEINAAKKAASNDSFGYT EDPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLV 420  
QY 421 KVVSWYDNEMSYTAQLVRTLEYFAKIAK 448  
Db 421 KVVSWYDNEMSYTAQLVRTLEYFAKIAK 448  
RESULT 3  
US-09-878-766A-12  
; Sequence 12, Application US/09878766A  
; Patent No. US20020044928A1  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus dysgalactiae  
US-09-878-766A-12

Query Match 72.7%; Score 1656.5; DB 9; Length 336;  
Best Local Similarity 79.6%; Pred. No. 3.4e-142;  
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;  
QY 28 MVVKVGINFGRIGRLAFRRIONVEGVEVTRINDLTD PNMHLAHLKDYDTTQGRFDGTVEV 87  
Db 1 MVVKVGINFGRIGRLAFRRIONVEGVEVTRINDLTD PNMHLAHLKDYDTTQGRFDGTVEV 60  
QY 88 KEGGFEVNGNFIKVS AERDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVS AE 147  
Db 61 KEGGFEVNGNFIKVS AERDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVS AE 81  
QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207  
Db 82 -----NIDWATDGV EIVLE 95  
QY 208 ATFFFAKKEAAEKHLHANGAKKVITAPG GNDVKT VVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFFAKKEAAEKHLHANGAKKVITAPG GNDVKT VVFNTHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILD GPHRGDLRRARAGAAANIVPNSTGAA 327  
Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILD GPHRGDLRRARAGAAANIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTG SVTELVTLDKNVSVDEINAAKKAASNDSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTG SVTELVTLDKNVSVDEINAAKKAASNDSFGYTE 275  
QY 388 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336

RESULT 4  
US-09-878-781-4

```
; Sequence 4, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-781-4

Query Match      72.7%; Score 1656.5; DB 10; Length 336;
Best Local Similarity 79.6%; Pred. No. 3.4e-142;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60

QY 88 KEGGFEVGNFIKVSARDPENIDWATDGEIVLEAEGTVEVKDGGFDVNGKFIKVSAAE 147
Db 61 KEGGFEVGNFIKVSARDPE-----81

QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 82 -----NIDWATDGEIVLE 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKKAASNDSFGYTE 275

QY 388 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448
Db 336 K 336

RESULT 5
US-10-650-369-12
; Sequence 12, Application US/10650369
; Publication No. US20040062774A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/10/650,369
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/878,766A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-10-650-369-12

Query Match      72.7%; Score 1656.5; DB 12; Length 336;
Best Local Similarity 79.6%; Pred. No. 3.4e-142;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60

QY 88 KEGGFEVGNFIKVSARDPENIDWATDGEIVLEAEGTVEVKDGGFDVNGKFIKVSAAE 147
Db 61 KEGGFEVGNFIKVSARDPE-----81

QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 82 -----NIDWATDGEIVLE 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKKAASNDSFGYTE 275

QY 388 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448
Db 336 K 336

RESULT 6
US-10-134-297-4
; Sequence 4, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055.20
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-10-134-297-4

Query Match      72.7%; Score 1656.5; DB 14; Length 336;
Best Local Similarity 79.6%; Pred. No. 3.4e-142;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
```



QY	88	KEGGFEVNGNFIKVS	AERDPENIDWATD	GVLEIVLEALEGT	VEVKDGGFDVNGKFIK	VS	AE	147	
Db	61	KEGGFEVNGNFIKVS	AERDPE	-----	-----	81			
QY	148	KDPEQIDWATD	GVLEIVLEIDGT	VEVKEGGFEVNG	QFVKVSAEREPA	NIDWATD	GVLEIVLE	207	
Db	82	-----	-----	-----	-----	NIDWATD	GVLEIVLE	95	
QY	208	ATSEFFAKKEAAE	KHLHANGAKK	VVITAPGGND	VKTVVFNTHD	ILDGTET	VISGASCTTN	267	
Db	96	ATGFFAKKEAAE	KHLHANGAKK	VVITAPGGND	VKTVVFNTHD	ILDGTET	VISGASCTTN	155	
QY	268	CLAPMAKALHDA	FGIOKGLMTTI	HAYTGDQ	MILDGPHRG	GLRRRAGA	ANIVPNSTGAA	327	
Db	156	CLAPMAKALHDA	FGIOKGLMTTI	HAYTGDQ	MILDGPHRG	GLRRRAGA	ANIVPNSTGAA	215	
QY	328	KAIGLVIPEL	NGKLDGAAQ	RVVPVPTGS	VTTELVTLD	KNVSVDE	INAAKKAAS	DSFGYTE	387
Db	216	KAIGLVIPEL	NGKLDGAAQ	RVVPVPTGS	VTTELVTLD	KNVSVDE	INAAKKAAS	DSFGYTE	275
QY	388	DPIVSSDIVG	SYGSLFDATQ	TKVMEVDG	SQVLKVVSWYD	NEMSYTAQL	VRTLEYF	AKIA	447
Db	276	DPIVSSDIVG	SYGSLFDATQ	TKVMEVDG	SQVLKVVSWYD	NEMSYTAQL	VRTLEYF	AKIA	335
QY	448	K	448						
Db	336	K	336						

## RESULT 7

```

US-10-282-122A-74379
; Sequence 74379, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614

```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74379
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74379

Query Match      72.7%;   Score 1655.5;   DB 12;   Length 336;
Best Local Similarity 79.3%;   Pred. No. 4.1e-142;
Matches 334;   Conservative 1;   Mismatches 1;   Indels 85;   Gaps 1;

QY      28  MVVKVGINGFRIGRLAPFRRIQNVEGVETRINDLTDNNMLAHLKDYDTTQGRFDGTVEV 87
      |||||||:|||||
Db      1  MVVKVGINGFRIGRLAPFRRIQNIEGVETRINDLTDNNMLAHLKDYDTTQGRFDGTVEV 60

QY      88  KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVVEKDGFDVNGKFIKVSAAE 147
      |||||||
Db      61  KEGGFEVNGNFIKVSARDPE----- 81

QY      148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207
      |||||||
Db      82  -----NIDWATDGVVEIVLE 95

QY      208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVISGASCTTN 267
      |||||||
Db      96  ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVISGASCTTN 155

QY      268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDPHRRGDLRRARAGAANI VPNSTGAA 327
      |||||||
Db      156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDPHRRGDLRRARAGAANI VPNSTGAA 215

QY      328 KAIGLVIPELNGKLDGAAQRVPVPTGSVTTELVTLLDKNVSVDEINAAKKAASNDSFGYTE 387
      |||||||
Db      216 KAIGLVIPELNGKLDGAAQRVPVPTGSVTTELVTLLDKNVSVDEINAAKKAASNDSFGYTE 275

QY      388 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSVSWYDNEMSYTAQLVRTLEVPKAKIA 447
      |||||||
Db      276 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSVSWYDNEMSYTAQLVRTLEVPKAKIA 335

QY      448 K 448
Db      336 K 336

```

## RESULT 8

```

US-09-878-781-14
; Sequence 14, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 14
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SpyGapC
; OTHER INFORMATION: protein
US-09-878-781-14

```

```

; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614

```





Db 276 DPIVSSDIIGMAYGSLFDATQTKVQTVDGNQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 11

US-09-878-781-8

; Sequence 8, Application US/09878781

; Publication No. US20030082781A1

; GENERAL INFORMATION:

; APPLICANT: Bolton, Alexandra J.

; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael

; APPLICANT: Potter, Andrew A.

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST

; TITLE OF INVENTION: STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0055

; CURRENT APPLICATION NUMBER: US/09/878,781

; CURRENT FILING DATE: 2002-09-10

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 336

; TYPE: PRT

; ORGANISM: Streptococcus uberis

US-09-878-781-8

Query Match 68.5%; Score 1559.5; DB 10; Length 336;

Best Local Similarity 73.9%; Pred. No. 2.3e-133;

Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVGINGGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87

Db 1 MVVKVGINGGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60

QY 88 KEGGFVNGNFIKVS AERDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVS AE 147

Db 61 KDGGFVNGNFIKVS AEK DPE----- 81

QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207

Db 82 -----NIDWATDGV EIVLE 95

QY 208 ATSFFAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDTGTTETVISGASCTTN 267

Db 96 ATGFFAKKAAAEKHLHANGAKKVITAPGDDVKTVVFNTHDILDTGTTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMILDGPHRGDLRRARAGAA NIVPNSTGAA 327

Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMILDGPHRGDLRRARAGAS NIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVTLDKNVSVDEINAAKAA S NDSFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVA VLEKETSV E E INAAKAA S NDSFGYTE 275

QY 388 DPIVSSDIVGSYGS LFDATQTKVMEVDGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 276 DPIVSSDIIGMAYGSLFDATQTKVQTVDGNQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 12

US-10-650-369-16

; Sequence 16, Application US/10650369

; Publication No. US20040062774A1

; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.

; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN

; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0057

; CURRENT APPLICATION NUMBER: US/10/650,369

; CURRENT FILING DATE: 2003-08-27

; PRIOR APPLICATION NUMBER: US/09/878,766A

; PRIOR FILING DATE: 2001-09-10

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 16

; LENGTH: 336

; TYPE: PRT

; ORGANISM: Streptococcus uberis

US-10-650-369-16

Query Match 68.5%; Score 1559.5; DB 12; Length 336;

Best Local Similarity 73.9%; Pred. No. 2.3e-133;

Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVGINGGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87

Db 1 MVVKVGINGGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60

QY 88 KEGGFVNGNFIKVS AERDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVS AE 147

Db 61 KDGGFVNGNFIKVS AEK DPE----- 81

QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207

Db 82 -----NIDWATDGV EIVLE 95

QY 208 ATSFFAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDTGTTETVISGASCTTN 267

Db 96 ATGFFAKKAAAEKHLHANGAKKVITAPGDDVKTVVFNTHDILDTGTTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMILDGPHRGDLRRARAGAA NIVPNSTGAA 327

Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMILDGPHRGDLRRARAGAS NIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVTLDKNVSVDEINAAKAA S NDSFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVA VLEKETSV E E INAAKAA S NDSFGYTE 275

QY 388 DPIVSSDIVGSYGS LFDATQTKVMEVDGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 276 DPIVSSDIIGMAYGSLFDATQTKVQTVDGNQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 13

US-10-134-297-8

; Sequence 8, Application US/10134297

; Publication No. US20030165524A1

; GENERAL INFORMATION:

; APPLICANT: Bolton, Alexandra J.

; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael

; APPLICANT: Potter, Andrew A.

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST

; TITLE OF INVENTION: STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0055.20

; CURRENT APPLICATION NUMBER: US/10/134,297

; CURRENT FILING DATE: 2002-04-26

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 336

; TYPE: PRT

; ORGANISM: Streptococcus uberis

US-10-134-297-8

Query Match 68.5%; Score 1559.5; DB 14; Length 336;  
Best Local Similarity 73.9%; Pred. No. 2.3e-133;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87  
|||||  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60  
|||||

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
|:|||||:|  
Db 61 KDGGEVNGNFIKVSAAEKDPE----- 81

QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207  
|||||  
Db 82 -----NIDWATDGVVEIVLE 95

QY 208 ATSEFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFTNHNHIDLDGTETVISGASCTTN 267  
|||  
Db 96 ATGFFAKKAAAEKHLHANGAKKVITAPGGDDVKTVVFTNHNHIDLDGTETVISGASCTTN 155  
|||

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 327  
|||  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 215  
|||

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVTLDKNVSVDEINAAKAAASNDSFGYTE 387  
|||  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVAVLEKETSVEEINAAKAAASNDSFGYTE 275  
|||

QY 388 DPIVSSDIVSVSGSLFDDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447  
|||  
Db 276 DPIVSSDIIGMAYGSLFDDATQTKVQTVDGNQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335  
|||

QY 448 K 448  
|  
Db 336 K 336

RESULT 14

US-09-878-766A-14  
; Sequence 14, Application US/09878766A  
; Patent No. US20020044928A1  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus agalactiae

US-09-878-766A-14

Query Match 68.4%; Score 1557.5; DB 9; Length 336;  
Best Local Similarity 73.9%; Pred. No. 3.5e-133;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87  
|||||  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRF----- 54  
|||||

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
|:|||||:|  
Db 55 ----- 54

QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207  
|||||  
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 95  
|||||

QY 208 ATSEFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFTNHNHIDLDGTETVISGASCTTN 267  
|||  
Db 96 ATGFFASKEKAGQHIHENGAKKVITAPGGNDVKTVVFTNHNHIDLDGTETVISGASCTTN 155  
|||

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 327  
|||||  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 215  
|||||

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVTLDKNVSVDEINAAKAAASNDSFGYTE 387  
|||||  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVALEKETSVEEINAAKAAASNDSFGYTE 275  
|||||

QY 388 DPIVSSDIVSVSGSLFDDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447  
|||||  
Db 276 DPIVSSDIIGMAYGSLFDDATQTKVQTVDGNQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335  
|||||

QY 448 K 448  
|  
Db 336 K 336

RESULT 14

US-09-878-766A-14  
; Sequence 14, Application US/09878766A  
; Patent No. US20020044928A1  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus agalactiae

US-09-878-766A-14

Query Match 68.4%; Score 1557.5; DB 9; Length 336;  
Best Local Similarity 73.9%; Pred. No. 3.5e-133;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87  
|||||  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRF----- 54  
|||||

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
|:|||||:|  
Db 55 ----- 54

QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207  
|||||  
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 95  
|||||

QY 208 ATSEFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFTNHNHIDLDGTETVISGASCTTN 267  
|||  
Db 96 ATGFFASKEKAGQHIHENGAKKVITAPGGNDVKTVVFTNHNHIDLDGTETVISGASCTTN 155  
|||

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 327  
|||||  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 215  
|||||

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVTLDKNVSVDEINAAKAAASNDSFGYTE 387  
|||||  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVALEKETSVEEINAAKAAASNDSFGYTE 275  
|||||

QY 388 DPIVSSDIVSVSGSLFDDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447  
|||||  
Db 276 DPIVSSDIIGMAYGSLFDDATQTKVQTVDGNQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335  
|||||

QY 448 K 448  
|  
Db 336 K 336

Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 95  
|||||

QY 208 ATSEFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFTNHNHIDLDGTETVISGASCTTN 267  
|||  
Db 96 ATGFFASKEKAGQHIHENGAKKVITAPGGNDVKTVVFTNHNHIDLDGTETVISGASCTTN 155  
|||

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 327  
|||||  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 215  
|||||

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVTLDKNVSVDEINAAKAAASNDSFGYTE 387  
|||||  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVALEKETSVEEINAAKAAASNDSFGYTE 275  
|||||

QY 388 DPIVSSDIVSVSGSLFDDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447  
|||||  
Db 276 DPIVSSDIIVGISGSLFDDATQTKVQTVDGNQLVKVSVWYDNEMSYTSQLVRLTLEYFAKIA 335  
|||||

QY 448 K 448  
|  
Db 336 K 336

RESULT 15

US-09-878-781-6  
; Sequence 6, Application US/09878781  
; Publication No. US20030082781A1  
; GENERAL INFORMATION:  
; APPLICANT: Bolton, Alexandra J.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; APPLICANT: Potter, Andrew A.  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST  
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0055  
; CURRENT APPLICATION NUMBER: US/09/878,781  
; CURRENT FILING DATE: 2002-09-10  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus agalactiae

US-09-878-781-6

Query Match 69.4%; Score 1557.5; DB 10; Length 336;  
Best Local Similarity 73.9%; Pred. No. 3.5e-133;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87  
|||||  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRF----- 54  
|||||

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
|:|||||:|  
Db 55 ----- 54

QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207  
|||||  
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 95  
|||||

QY 208 ATSEFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFTNHNHIDLDGTETVISGASCTTN 267  
|||  
Db 96 ATGFFASKEKAGQHIHENGAKKVITAPGGNDVKTVVFTNHNHIDLDGTETVISGASCTTN 155  
|||

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 327  
|||||  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 215  
|||||

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVTLDKNVSVDEINAAKAAASNDSFGYTE 387  
|||||  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTSVTTELVALEKETSVEEINAAKAAASNDSFGYTE 275  
|||||

QY	388	DPIVSSDIVGVSGLF	DATQTKVMEVDGSQLVKVVS	WYDNEMSYTAQLV	RTLEYFAKIA	447
Db	276	DPIVSSDIVGVSGLF	DATQTKVQTV	DGNQLVKVVS	WYDNEMSYTSQLV	RTLEYFAKIA 335
QY	448	K	448			
Db	336	K	336			

Search completed: September 15, 2004, 10:22:15  
Job time : 133 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:20:04 ; Search time 41 Seconds  
(without alignments)  
1051.068 Million cell updates/sec

Title: US-10-650-369-22  
Perfect score: 2278  
Sequence: 1 MKKITGRIILLAVIILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1652.5	72.5	336	2 A42963	glyceraldehyde-3-p
2	1558.5	68.4	336	2 S71350	glyceraldehyde-3-p
3	1535	67.4	335	2 F95235	glyceraldehyde 3-p
4	1535	67.4	359	2 G98099	glyceraldehyde-3-p
5	1357.5	59.6	336	2 F86905	hypothetical prote
6	1280	56.2	337	2 G86694	hypothetical prote
7	1234	54.2	334	2 S34254	glyceraldehyde-3-p
8	1205.5	52.9	336	2 AC1382	glyceraldehyde 3-p
9	1204.5	52.9	336	2 AD1751	glyceraldehyde 3-p
10	1203	52.8	334	2 C96987	glyceraldehyde 3-p
11	1145	50.3	334	2 B82019	glyceraldehyde-3-p
12	1132.5	49.7	334	2 E81001	glyceraldehyde-3-p
13	1132	49.7	334	2 E81001	glyceraldehyde 3-p
14	955.5	41.9	338	2 T09633	glyceraldehyde-3-p
15	941.5	41.3	333	2 F90881	glyceraldehyde 3-p
16	941.5	41.3	333	2 C85737	glyceraldehyde-3-p
17	870.5	38.2	337	2 S73737	glyceraldehyde-3-p
18	861.5	37.8	349	2 F90517	glyceraldehyde 3-p
19	854.5	37.5	337	2 C64233	glyceraldehyde-3-p
20	851	37.4	336	2 A43260	glyceraldehyde-3-p
21	849	37.3	335	1 DEBSG	glyceraldehyde-3-p
22	845	37.1	335	2 S12696	glyceraldehyde-3-p
23	836	36.7	335	2 H84094	glyceraldehyde-3-p
24	830	36.4	335	1 DEBSGF	glyceraldehyde-3-p
25	820.5	36.0	335	2 A70107	probable glycerald
26	819.5	36.0	333	1 DEHGGT	glyceraldehyde-3-p
27	793.5	34.8	342	2 F70391	glyceraldehyde-3-p
28	790	34.7	336	2 T36020	glyceraldehyde-3-p
29	786.5	34.5	339	2 G70915	glyceraldehyde-3-p

30	783.5	34.4	339	2 S72763	glyceraldehyde-3-p
31	781.5	34.3	334	2 AI0262	glyceraldehyde-3-p
32	780.5	34.3	331	2 F82131	glyceraldehyde 3-p
33	769.5	33.8	330	2 E75408	glyceraldehyde 3-p
34	768.5	33.7	331	2 AG0711	glyceraldehyde 3-p
35	764.5	33.6	331	1 DEECG3	glyceraldehyde-3-p
36	764.5	33.6	331	2 H90939	glyceraldehyde-3-p
37	764.5	33.6	331	2 D85788	glyceraldehyde-3-p
38	754	33.1	339	2 G64041	glyceraldehyde-3-p
39	748.5	32.9	331	1 DEUTC	glyceraldehyde-3-p
40	739	32.4	332	2 JC4373	glyceraldehyde-3-p
41	734	32.2	336	2 T40235	glyceraldehyde-3-p
42	732.5	32.2	337	2 AF2438	glyceraldehyde-3-p
43	730.5	32.1	337	2 S54141	glyceraldehyde-3-p
44	729	32.0	337	2 S40610	glyceraldehyde-3-p
45	727.5	31.9	344	2 E84043	glyceraldehyde-3-p

ALIGNMENTS

RESULT 1  
A42963  
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus  
N;Alternate names: plasmin receptor  
C;Species: Streptococcus sp.  
C;Date: 10-Mar-1994 #sequence revision 10-Mar-1994 #text\_change 03-Jun-2002  
C;Accession: A42963; B42963; JH0750  
R;Lottenberg, R.; Broder, C.C.; Boyle, M.D.; Kain, S.J.; Schroeder, B.L.; Curtiss III, R.  
J. Bacteriol. 174, 5204-5210, 1992  
A;Title: Cloning, sequence analysis, and expression in Escherichia coli of a streptococcal  
A;Reference number: A42963; MUID:92355491; PMID:1322883  
A;Accession: A42963  
A;Molecule type: DNA  
A;Residues: 1-336 <LOT>  
A;Experimental source: group A, strain 64/14  
A;Note: sequence extracted from NCBI backbone (NCBIP:110308)  
A;Accession: B42963  
A;Molecule type: protein  
A;Residues: 2-74;161-164,'X',166-174;187-211,'X',213-217 <LO2>  
R;Pancholi, V.; Fischetti, V.A.  
J. Exp. Med. 176, 415-426, 1992  
A;Title: A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate  
A;Reference number: JH0750; MUID:92364544; PMID:1500854  
A;Accession: JH0750  
A;Molecule type: protein  
A;Residues: 2-30,'A',32-40 <PAN>  
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase  
C;Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase  
F;152/Active site: Cys #status predicted

Query Match	72.5%	Score	1652.5;	DB	2;	Length	336;
Best Local Similarity	79.1%;	Pred.	No. 9.1e-101;				
Matches	333;	Conservative	2;	Mismatches	1;	Indels	85;
Gaps							1;
QY	28	MVVKVINGFGRIGRLAFRRIQNVGEVETRINDLTDPNNMLAHLKYDTTQGRFDGTVEV	87				
Db	1	MVVKVINGFGRIGRLAFRRIQNVGEVETRINDLTDPNNMLAHLKYDTTQGRFDGTVEV	60				
QY	88	KEGGFEVGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGFWDWNGKFIKVS	147				
Db	61	KEGGFEVGNFIKVSARDPE	81				
QY	148	KOPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPA	207				
Db	82	NIDWATDGVLE	95				
QY	208	ATSFPAKKEAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTETVISGASCTTN	267				
Db	96	ATGFFAKKEAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTETVISGASCTTN	155				
QY	268	CLAPMAKALHDAFGIQKGLMTTTHAYTGDMILDGPHRGDGLRRRAGAANIVPNSTGAA	327				

Db 156 CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMILDGPHRGDDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQORVPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQORVPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFGYTE 275

QY 388 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 276 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 2

S71350

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus

C;Species: Streptococcus "equisimilis"

C;Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 03-Jun-2002

C;Accession: S71350

R;Gase, K.; Gase, A.; Schirmer, H.; Malke, H.

Eur. J. Biochem. 239, 42-51, 1996

A;Title: Cloning, sequencing and functional overexpression of the Streptococcus equisimilis protein. Purification and biochemical characterization of the protein.

A;Reference number: S71350; MUID:96305364; PMID:8706717

A;Accession: S71350

A;Molecule type: DNA

A;Residues: 1-336 <GAS>

A;Cross-references: EMBL:X97788; NID:g1478268; PIDN:CAA66377.1; PID:g1478269

A;Experimental source: strain H46A

C;Genetics:

A;Gene: gapC

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C;Keywords: Gluconeogenesis; glycolysis; NAD; oxidoreductase

F;4-34/Region: beta-alpha-beta NAD nucleotide-binding fold

F;152/Active site: Cys #status predicted

Query Match 68.4%; Score 1558.5; DB 2; Length 336;

Best Local Similarity 76.0%; Pred. No. 1.3e-94;

Matches 320; Conservative 2; Mismatches 14; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 87

Db 1 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147

Db 61 KEGGFEVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVLE 207

Db 82 -----NIDWATDGVLE 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267

Db 96 ATGFFAKKEAAEKPLHANGAKKVITAPGGNDVKQLFSTLTTSILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMILDGPHRGDDLRRARAGAAANIVPNSTGAA 327

Db 156 CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMIVDGHRRGGDLRRARAGAAANIVPNSTGAR 215

QY 328 KAIGLVIPELNGKLDGAAQORVPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQORVPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFGYTE 275

QY 388 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 276 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 3

F95235

glyceraldehyde 3-phosphate dehydrogenase [imported] - Streptococcus pneumoniae (strain

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001

C;Accession: F95235

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; He

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap

ple, nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: F95235

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-335 <KUR>

A;Cross-references: GB:AE005672; PIDN:AAK76079.1; PID:g14973522; GSPDB:GN00164; TIGR:SP

A;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP2012

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 67.4%; Score 1535; DB 2; Length 335;

Best Local Similarity 73.6%; Pred. No. 4.4e-93;

Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 87

Db 1 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147

Db 61 KEGGFEVNGKFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVLE 207

Db 82 -----QIDWATDGVLE 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267

Db 96 ATGFFAKKEAAEKHL-KGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 154

QY 268 CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMILDGPHRGDDLRRARAGAAANIVPNSTGAA 327

Db 155 CLAPMAKALQDNFVGVVEGLMTTTHAYTGDQMILDGPHRGDDLRRARAGAAANIVPNSTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQORVPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFGYTE 387

Db 215 KAIGLVIPELNGKLDGSAQORVPTGVSVELVAVLEKNVTDEVNAAKKAASNESYGYTE 274

QY 388 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 275 DPIVSSDIVGMSYGSFLFDATQTKVLVDVGKQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 334

QY 448 K 448

Db 335 K 335

RESULT 4

G98099

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) [imported] - St

C;Species: Streptococcus pneumoniae

C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 03-Jun-2002

C;Accession: G98099

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A;Reference number: A97872; MUID:21429245; PMID:11544234

A;Accession: G98099

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-359 <KUR>

A;Cross-references: GB:AE007317; PIDN:AAL00628.1; PID:g15459513; GSPDB:GN00174

C;Genetics:

A;Gene: gapA

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C;Keywords: oxidoreductase

Query Match 67.4%; Score 1535; DB 2; Length 359;  
Best Local Similarity 73.6%; Pred. No. 4.8e-93;  
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDENMLAHLKYDTTQGRFDGTVEV 87  
Db 25 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDVPMLAHLKYDTTQGRFDGTVEV 84

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVS AE 147  
Db 85 KEGGFEVNGKFIKVSARDPE----- 105

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207  
Db 106 -----QIDWATDGVIEVLE 119

QY 208 ATFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 120 ATGFFAKKEAAEKHL-KGAKKVVITAPGGNDVKTVVFNTHDVLDTGTETVISGASCTTN 178

QY 268 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 327  
Db 179 CLAPMAKALQDNFVVEGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 238

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTTELVTLDKNVSVDEINAAKMAASNSFGYTE 387  
Db 239 KAIGLVIPELNGKLDGSAQRPVPTGSGVTTELVALEKNVTVDENVAAKMAASNSFGYTE 298

QY 388 DPIVSSDIVGVSYSLSFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447  
Db 299 DPIVSSDIVGMSYLSFDATQTKVLVDVGKQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 358

QY 448 K 448

Db 359 K 359

RESULT 5  
F86905  
hypothetical protein gapB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C;Species: Lactococcus lactis subsp. lactis  
C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C;Accession: F86905  
R;Bolotin, A.; Wincker, P.; Mauder, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich; Genome Res. 11, 731-753, 2001  
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss  
A;Reference number: A86625; MUID:21235186; PMID:11337471  
A;Accession: F86905  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-336 <STO>  
A;Cross-references: GB:AE005176; PID:g12725315; PIDN:AAK06344.1; GSPDB:GN00146  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: gapB  
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 59.6%; Score 1357.5; DB 2; Length 336;  
Best Local Similarity 65.1%; Pred. No. 1.8e-81;  
Matches 274; Conservative 21; Mismatches 41; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDENMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDVPMLAHLKYDTTQGRFDGTVEV 54

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVS AE 147  
Db 55 ----- 54

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207  
Db 55 -----DGKVEVKDGGFEVNGKFKVTAESNPANINWAEVGAIEVLE 95

QY 208 ATFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFATKEKAEQHLHANGAKKVVITAPGSDVKTIIVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 327  
Db 156 CLAPMADTLNKQFGIKVGTMTTVHGYTGDQMTLDGPHRGDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTTELVTLDKNVSVDEINAAKMAASNSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGHAQRPVPTGSLTELVTILNKEVTVDENVAAKMAASNSFGYNE 275

QY 388 DPIVSSDIVGVSYSLSFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447  
Db 276 DPIVSSDIVGINSLSFDATQTEVTSADGAQLVKTVSVWYDNEMSYTNLVRTLAYFAKIA 335

QY 448 K 448  
Db 336 K 336

RESULT 6  
G86694  
hypothetical protein gapA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C;Species: Lactococcus lactis subsp. lactis  
C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C;Accession: G86694  
R;Bolotin, A.; Wincker, P.; Mauder, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich; Genome Res. 11, 731-753, 2001  
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss  
A;Reference number: A86625; MUID:21235186; PMID:11337471  
A;Accession: G86694  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-337 <STO>  
A;Cross-references: GB:AE005176; PID:g12723446; PIDN:AAK04657.1; GSPDB:GN00146  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: gapA  
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 56.2%; Score 1280; DB 2; Length 337;  
Best Local Similarity 61.4%; Pred. No. 2e-76;  
Matches 259; Conservative 31; Mismatches 46; Indels 86; Gaps 2;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDENMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLALRRIOEVEGEVAHINDLTDVPMLAHLKYDTTQGRFDGTVEV 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVS AE 147  
Db 61 KEDGFDVNGKFKVTAERNPE----- 81

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207  
Db 82 -----DIQWADSGVEIVLE 95

QY 208 ATFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFATKEKAEKHLHPGGAKKVILITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155











Search completed: September 15, 2004, 10:27:49  
Job time : 43 secs

QY	31	KVGINFGFRIGRLAFERRIQNVE-GVEVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEVKE	89
Db	3	KVGINFGFRIGRLVLRRLLEVKSNIADVAINDLTSPKILAYLLKXHSNYGPF-----	54
QY	90	GGFEVNGNFIKVSARDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVSAEKD	149
Db	55	-----	54
QY	150	PEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLEAT	209
Db	55	---PWSVDFTEDSLIVD-----GKSIAYVAEKEAKNIPWKAGAEIIVECT	97
QY	210	SFFAKKEAAEKHLHANGAKKVVITAPGNGDVKTVVFNTHDILDGTETVISGASCTTNCL	269

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:11:48 ; Search time 25 Seconds  
(without alignments)  
933.097 Million cell updates/sec

Title: US-10-650-369-22  
Perfect score: 2278  
Sequence: 1 MKKITGIILLLLAVIILSAC.....EMSYTAQLVRLTLEYFAKIAK 448

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues 141681  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1650.5	72.5	335	1 G3P_STRPY	P50467 streptococc
2	1646.5	72.3	335	1 G3P_STRP3	Q8k8m9 streptococc
3	1553.5	68.2	335	1 G3P_STREQ	Q59906 streptococc
4	1280	56.2	337	1 G3P_LACLA	P52987 lactococcus
5	1234	54.2	334	1 G3P_CLOPA	Q59309 clostridium
6	1203	52.8	334	1 G3P_CLOAB	O52631 clostridium
7	1138.5	50.0	336	1 G3P1_STAEP	Q8cpx5 staphylococ
8	1132.5	49.7	336	1 G3P1_STAAM	Q9z5c5 staphylococ
9	955.5	41.9	338	1 G3P_LACDE	O32755 lactobacill
10	941.5	41.3	333	1 G3P3_ECO57	P58072 escherichia
11	940.5	41.3	333	1 G3P3_ECOLI	P33898 escherichia
12	870.5	38.2	337	1 G3P_MYCPN	P75358 mycoplasma
13	854.5	37.5	337	1 G3P_MYCGE	P47543 mycoplasma
14	850	37.3	334	1 G3P_CORGL	Q01651 corynebacte
15	844	37.1	334	1 G3P1_BACSU	P09124 bacillus su
16	840	36.9	334	1 G3P_BACME	P23722 bacillus me
17	825	36.2	334	1 G3P_BACST	P00362 bacillus st
18	820.5	36.0	335	1 G3P_BORBU	P46795 borrelia bu
19	818.5	35.9	332	1 G3P_THEMA	P17721 thermotoga
20	793.5	34.8	342	1 G3P_AQUAE	O67161 aquifex aeo
21	790	34.7	336	1 G3P_STRCO	Q9z5l8 streptomyce
22	786.5	34.5	339	1 G3P_MYCTU	O06822 mycobacteri
23	783.5	34.4	339	1 G3P_MYCLE	P46713 mycobacteri
24	777.5	34.1	339	1 G3P_MYCAV	P94915 mycobacteri
25	763.5	33.5	330	1 G3P1_SALTY	P24165 salmonella
26	759.5	33.3	330	1 G3P1_ECOLI	P06977 escherichia
27	754	33.1	339	1 G3P_HAEIN	P44304 haemophilus
28	743.5	32.6	330	1 G3PC_TRYBB	P10097 trypanosoma
29	740	32.5	332	1 G3P_RALSO	P52694 ralstonia s
30	739	32.4	332	1 G3P_STRAU	Q59800 streptomyce
31	737	32.4	337	1 G3P1_RHIRA	Q9c136 rhizomucor
32	737	32.4	337	1 G3P_MONAN	P53430 monascus an
33	734	32.2	336	1 G3P1_SCHPO	P78958 schizosacch

34	732.5	32.2	337	1 G3P2_ANASP	P58554 anabaena sp
35	729	32.0	337	1 G3P_GLAPU	Q00584 claviceps p
36	725	31.8	331	1 G3P_THEAQ	P00361 thermus aqu
37	724.5	31.8	336	1 G3P2_SYNY3	P80505 synecocyst
38	723	31.7	338	1 G3P_NEUCR	P54118 neurospora
39	722	31.7	337	1 G3P_COCHE	P29497 cochliobolu
40	721	31.7	337	1 G3P_CURLU	P28844 curvularia
41	719.5	31.6	333	1 G3P_STRAE	P54226 streptomyce
42	717.5	31.5	330	1 G3PC_LEIME	Q01558 leishmania
43	717	31.5	335	1 G3P2_SCHPO	O43026 schizosacch
44	715	31.4	336	1 G3P_ASPNG	Q12552 aspergillus
45	713	31.3	336	1 G3P_EMENI	P20445 emericella

ALIGNMENTS

RESULT 1  
G3P\_STRPY  
ID\_G3P\_STRPY STANDARD; PRT; 335 AA.  
AC P50467;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)  
DE (Plasminogen-binding protein) (Plasmin receptor).  
GN GAP OR PLR OR GAPA OR SPY0274 OR SPYM18\_0261.  
OS Streptococcus pyogenes, and  
OS Streptococcus pyogenes (serotype M18).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314, 186103;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-73; 160-173 AND 186-216.  
RC STRAIN=64/14;  
RX MEDLINE=23355491; PubMed=1322883;  
RA Lottenberg R., Broder C.C., Boyle M.D., Kain S.J., Schroeder B.L.,  
RA Curtiss R. III;  
RA "Cloning, sequence analysis, and expression in Escherichia coli of a  
RT streptococcal plasmin receptor.";  
RL J. Bacteriol. 174:5204-5210(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;  
RX MEDLINE=21192684; PubMed=11296296;  
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
RT "Complete genome sequence of an M1 strain of Streptococcus  
pyogenes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MGAS8232 / Serotype M18;  
RX MEDLINE=21927593; PubMed=11917108;  
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,  
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,  
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,  
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;  
RT "Genome sequence and comparative microarray analysis of serotype M18  
RT group A Streptococcus strains associated with acute rheumatic fever  
RT outbreaks.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).  
RN [4]  
RP SEQUENCE OF 1-11; 20-30; 103-128; 162-171 AND 199-215.  
RC STRAIN=JRS4 / Serotype M6;  
RA Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,  
RA VanBogelen R.A.;  
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes  
RT proteins.";  
RL Submitted (MAY-2000) to Swiss-Prot.  
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.



Db 1 VVKVINGFGRIGRLAFRRRIQIEGVETVRINDLTDPNMLAHLKYDITQGRFDGTVEVK 60

QY 89 EGGFEVNGNFIVKSAERDPENIDWATDGVVEIVLEALEGIVEVKGDFVNGKFIKVSAAK 148

Db 61 EGGFEVNGNFIVKSAERDPE----- 80

QY 149 DPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEA 208

Db 81 -----NIDWATDGVVEIVLEA 95

QY 209 TSFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTHDILDTETVISGASCTTNC 268

Db 96 TGFFAKKEAAEKHLHTNGAKKVITAPGGNDVKTVVFNTHDILDTETVISGASCTTNC 155

QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAAK 328

Db 156 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAAK 215

QY 329 AIGLVIPELNGKLDGAAQRPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFGYTED 388

Db 216 AIGLVIPELNGKLDGAAQRPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFGYTED 275

QY 389 PIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 448

Db 276 PIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 335

RESULT 3

G3P\_STREQ

ID G3P STREQ STANDARD; PRT; 335 AA.

AC Q59306;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)

DE (Plasminogen-binding protein) (Plasmin receptor).

GN GAP OR GAPC.

OS Streptococcus equisimilis.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=119602;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H46A;

RX MEDLINE=96305364; PubMed=8706717;

RA Gase K., Gase A., Schirmer H., Malke H.;

RT "Cloning, sequencing and functional overexpression of the

RT Streptococcus equisimilis H46A gapC gene encoding a

RT glyceraldehyde-3-phosphate dehydrogenase that also functions as a

RT plasmin(ogen)-binding protein. Purification and biochemical

RT characterization of the protein.";

RL Eur. J. Biochem. 239:42-51(1996).

CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.

CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +

CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -!- PATHWAY: Second phase of glycolysis; first step.

CC -!- SUBUNIT: Homotetramer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate

CC dehydrogenase family.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

CC EMBL; X97788; CAA66377.1; -.

CC EMBL; Y12602; CAA73174.1; -.

CC PIR; S71350; S71350.

CC HSSP; P00362; IGD1.

DR InterPro; IPR000173; GAP\_dhdrogenase.

DR InterPro; IPR006424; GAPDH-I.

DR Pfam; PF00044; gpdh; 1.

DR Pfam; PF02800; gpdh\_C; 1.

DR PRINTS; PR00078; G3PDHDRGNASE.

DR TIGRFAMS; TIGR01534; GAPDH-I; 1.

DR PROSITE; PS00071; GAPDH; 1.

KW Glycolysis; Oxidoreductase; NAD.

FT INIT\_MET 0

FT BINDING 151 151

FT BY SIMILARITY.

FT GLYCERALDEHYDE 3-PHOSPHATE (BY

FT SIMILARITY).

FT ACT\_SITE 178 178

FT ACTIVATES THIOL GROUP DURING CATALYSIS

FT (BY SIMILARITY).

SQ SEQUENCE 335 AA; 35739 MW; FE7ACFD7663E46 CRC64;

Query Match 68.2%; Score 1553.5; DB 1; Length 335;

Best Local Similarity 76.0%; Pred. No. 9.5e-88;

Matches 319; Conservative 2; Mismatches 14; Indels 85; Gaps 1;

QY 29 VVKVINGFGRIGRLAFRRRIQNVESVEVTRINDLTDPNMLAHLKYDITQGRFDGTVEVK 88

Db 1 VVKVINGFGRIGRLAFRRRIQNVESVEVTRINDLTDPNMLAHLKYDITQGRFDGTVEVK 60

QY 89 EGGFEVNGNFIVKSAERDPENIDWATDGVVEIVLEALEGTVEVKGDFVNGKFIKVSAAK 148

Db 61 EGGFEVNGNFIVKSAERDPE----- 80

QY 149 DPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEA 208

Db 81 -----NIDWATDGVVEIVLEA 95

QY 209 TSFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTHDILDTETVISGASCTTNC 268

Db 96 TGFFAKKEAAEKPLHANGAKKVITAPGGNDVQKLFSTLTTSILDGTETVISGASCTTNC 155

QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAAK 328

Db 156 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMIVDGHRGGDLRRARAGAAANIVPNSTGARK 215

QY 329 AIGLVIPELNGKLDGAAQRPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFGYTED 388

Db 216 AIGLVIPELNGKLDGAAQRPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFGYTED 275

QY 389 PIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 448

Db 276 PIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 335

RESULT 4

G3P\_LACLA

ID G3P LACLA STANDARD; PRT; 337 AA.

AC P52987;

DT 01-OCT-1996 (Rel. 34, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).

GN GAP OR LL0559.

OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

OX NCBI\_TaxID=1360;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LMO230;

RX MEDLINE=95291425; PubMed=7773380;

RA Cancilla M.R., Hillier A.J., Davidson B.E.;

RT "Lactococcus lactis glyceraldehyde-3-phosphate dehydrogenase gene,

RT gap: further evidence for strongly biased codon usage in glycolytic

RT pathway genes.";

RL Microbiology 141:1027-1036(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=IL1403;

RX MEDLINE=21235186; PubMed=11337471;







QY 149 DPEQIDWATDGVVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEA 208  
Db 85 -----GEL-----GIDVVEVC 95  
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDILDTGTTTISGASCTTNC 268  
Db 96 TGFTTKKEAEAHVRA-GAKKVVISAPAGNDLKTIVFNVNNDLDTGTTTISGASCTTNC 154  
QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDMILDGPHRGDGLRRARAGANIVPNSTGA 328  
Db 155 LAPMAKVLNDKFGIEKGFMTTIHAYTNDQNTLDGPHRGDGLRRARAGANIVPNSTGA 214  
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVTLVTLVTLVTLVTLVTLVTLVTLV 388  
Db 215 AIAQVIPLEKGLDGNAGQRPVPTGVTGLVTLVTLVTLVTLVTLVTLVTLVTLVTLV 274  
QY 389 PIVSSDIVGVSIGSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 448  
Db 275 EIVSADVVGISGSLFDATLTIKIVDVGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 334

RESULT 6

G3P\_CLOAB STANDARD; PRT; 334 AA.  
AC O52631;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).  
GN GAP OR GAPC OR CAC0709.  
OS Clostridium acetobutylicum.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1488;  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
RX MEDLINE=99392446; PubMed=10463150;  
RA Schreiber W., Durre P.;  
RT "The glycerolaldehyde-3-phosphate dehydrogenase of Clostridium  
RT acetobutylicum: isolation and purification of the enzyme, and  
RT sequencing and localization of the gap gene within a cluster of other  
RT glycolytic genes."  
RL Microbiology 145:1839-1847(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
RX MEDLINE=21359325; PubMed=11466286;  
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
RA Bennett G.N., Koonin E.V., Smith D.R.;  
RT "Genome sequence and comparative analysis of the solvent-producing  
RT bacterium Clostridium acetobutylicum."  
RL J. Bacteriol. 183:4823-4838(2001).  
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
CC -!- PATHWAY: Second phase of glycolysis; first step.  
CC -!- SUBUNIT: Homotetramer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
CC dehydrogenase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF043386; AAC13160.1; -.  
CC EMBL; AE007586; AAK78686.1; -.

DR PIR; C96987; C96987.  
DR HSSP; P17721; 1HDG.  
DR InterPro; IPR000173; GAP dhhydrogenase.  
DR InterPro; IPR006424; GAPDH-I.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh C; 1.  
DR PRINTS; PR00078; G3PDHGRNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.  
FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE.  
FT ACT SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.  
SQ SEQUENCE 334 AA; 35850 MW; 10C52A174BE789B5 CRC64;  
Query Match 52.8%; Score 1203; DB 1; Length 334;  
Best Local Similarity 57.4%; Pred. No. 2e-66;  
Matches 241; Conservative 41; Mismatches 52; Indels 86; Gaps 3;  
QY 29 VVKVINGFGRIGRLAFRRIQNVGEVETRIINDLTDPNMLAHLKYDTTQGRFDGTVEVK 88  
Db 1 MAKIAINGFGRIGRLALRRILEVPGLEVVAINDLTDKMLAHLKYDSSQGRFNGEIEVK 60  
QY 89 EGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVS 148  
Db 61 EGAFVNGKEVKVFAEADPEKLPWG-----DL----- 87  
QY 149 DPEQIDWATDGVVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEA 208  
Db 88 -----GIDVVEVC 95  
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDILDTGTTTISGASCTTNC 268  
Db 96 TGFTTKKEAEAHVRA-GAKKVVISAPAGNDLKTIVFNVNNDLDTGTTTISGASCTTNC 154  
QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDMILDGPHRGDGLRRARAGANIVPNSTGA 328  
Db 155 LAPMAKVLNDKFGIEKGFMTTIHAYTNDQNTLDGPHRGDGLRRARAGANIVPNSTGA 214  
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVTLVTLVTLVTLVTLVTLVTLVTLV 388  
Db 215 AISOVIPDLAGKLDGNAGQRPVPTGVTGLVTLVTLVTLVTLVTLVTLVTLVTLVTLV 274  
QY 389 PIVSSDIVGVSIGSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 448  
Db 275 PIVSADVVGISGSLFDATLTIKIVDVGSQLVKVTAAYDNEMSYTAQLVRLTLEYFAK 334

RESULT 7

G3P1\_STAEP STANDARD; PRT; 336 AA.  
AC Q8CPY5;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH 1).  
GN GAPC OR GAP OR SE0557.  
OS Staphylococcus epidermidis.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 12228;  
RX PubMed=12950922;  
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,  
RA Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,  
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;  
RT "Genome-based analysis of virulence genes in a non-biofilm-forming  
RT Staphylococcus epidermidis strain (ATCC 12228).";  
RL Mol. Microbiol. 49:1577-1593(2003).  
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
CC -!- PATHWAY: Second phase of glycolysis; first step.  
CC -!- SUBUNIT: Homotetramer (By similarity).











RL J. Bacteriol. 173:6118-6123(1991).  
RN [6]  
RP SEQUENCE OF 1-13 FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=88232416; PubMed=2836696;  
RA Nakamura H., Murakami H., Yamato I., Anraku Y.;  
RT "Nucleotide sequence of the cybB gene encoding cytochrome b561 in  
RT Escherichia coli K12";  
RL Mol. Gen. Genet. 212:1-5(1988).  
RN [7]  
RP SEQUENCE OF 1-165 FROM N.A.  
RA Krawetz S.A.;  
RL Submitted (OCT-1990) to the EMBL/GenBank/DDBJ databases.  
RN [8]  
RP SEQUENCE OF 1-52 FROM N.A.  
RX MEDLINE=87053179; PubMed=3780374;  
RA Krawetz S.A., Connor W., Cannon P.D., Dixon G.H.;  
RT "A vector-primer-cloner-sequencer plasmid for the construction of cDNA  
RT libraries: evidence for a rat glyceraldehyde-3-phosphate  
RT dehydrogenase-like mRNA and a ferritin mRNA within testis.";  
RL DNA 5:427-435(1986).  
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
CC -!- PATHWAY: Second phase of glycolysis; first step.  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
CC dehydrogenase family.  
CC -!- CAUTION: IN THE K12 STRAIN THIS GENE IS DISRUPTED BY A STOP CODON  
CC AND A FRAMESHIFT. IT SEEMS TO BE INTACT IN A NUMBER OF WILD  
CC STRAINS.  
CC -!- CAUTION: Ref.7 and Ref.8 sequences were originally thought to  
CC originate from rat.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L09067; AAA23856.1; ALT\_FRAME.  
DR EMBL; AE000239; AAC74498.1; ALT\_FRAME.  
DR EMBL; AE000239; AAC74499.1; ALT\_FRAME.  
DR EMBL; D90780; BAA15033.1; ALT\_FRAME.  
DR EMBL; D90781; BAA15038.1; ALT\_FRAME.  
DR EMBL; M64541; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X07569; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X54798; CAA38569.1; -.  
DR EMBL; M14166; AAA41178.1; -.  
DR HSSP; P17721; 1HDG.  
DR EcoGene; EG12103; gapC.  
DR InterPro; IPR000173; GAP\_dhdrogenase.  
DR InterPro; IPR006424; GAPDH-I.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh.C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.  
FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE  
FT (BY SIMILARITY).  
FT ACT\_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS  
FT (BY SIMILARITY).  
FT CONFLICT 39 39 K -> Y (IN REF. 7 AND 8).  
SQ SEQUENCE 333 AA; 35649 MW; E8822329737376B0A0 CRC64;  
  
Query Match 41.3%; Score 940.5; DB 1; Length 333;  
Best Local Similarity 47.5%; Pred. No. 1.9e-50;  
Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;  
  
QY 31 KVGINGFGRIGRLAFRRIQNVE-GVEVTRINDLDPNMLAHLKYDTTQGRFDGTVEVKE 89

Db 3 KVGINGFGRIGRLVLRLLLEVKSNIDVVAINDLTSPKILAYLLKHDSNYGPF----- 54  
QY 90 GGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAEKD 149  
Db 55 ----- 54  
QY 150 PEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEAT 209  
Db 55 ----PWSVDFTEDSLIVD-----GKSIAYAEKEAKNIPWKAKGAEIIVECT 97  
QY 210 SFFAKKEAAEKHLHANGAKKVITAPGCGNDVKTVVFNTHDILDGTEIVISGASCTTNCL 269  
Db 98 GFYTSAEKSAHLDA-GAKKVLISAPAG-EMKTIYVNVNDDTLDNGDNTIVSVASCTTNCL 155  
QY 270 APMKALHDAFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAAGA 329  
Db 156 APMKALHDSFGIEVGTMTTIHAYTGTQSLVDGP-RGKDLRASRAAAENIIPHTTGAAGA 214  
QY 330 IGLVIPLELNGKLDGAAQRPVPTGVSVELVVTLDKNVSVDEINAAKAA--SNDSFGYTE 387  
Db 215 IGLVIPLESGKLKGAHQRPVKTSVELVSVILGKKVTAEVNNALKQATTNNESEFGYTD 274  
QY 388 DPIVSSDIVGSYGSFLFDATQTQKVMEDVDSQLVKVSWYDNEMSYTAQLVRTLLEYFAKI 446  
Db 275 EEIVSSDIIGSHFGSVFDATQTEITAVGDLQLVKTVAWYDNEYGFVTQIRTLEKFAKL 333  
  
RESULT 12  
G3P\_MYCPN STANDARD; PRT; 337 AA.  
ID G3P\_MYCPN  
AC P75358;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).  
GN GAPA OR GAP OR MPN430 OR MP411.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
RA Herrmann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
RT pneumoniae";  
RL Nucleic Acids Res. 24:4420-4449(1996).  
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
CC -!- PATHWAY: Second phase of glycolysis; first step.  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
CC dehydrogenase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE000040; AAB96059.1; -.  
DR PIR; S73737; S73737.  
DR HSSP; P17721; 1HDG.  
DR InterPro; IPR000173; GAP\_dhdrogenase.  
DR InterPro; IPR006424; GAPDH-I.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh.C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.

DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.  
FT BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE  
FT ACT\_SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS  
FT ACT\_SITE 184 184 (BY SIMILARITY).  
SQ SEQUENCE 337 AA; 36805 MW; 550747A529ABCA83 CRC64;  
  
Query Match 38.2%; Score 870.5; DB 1; Length 337;  
Best Local Similarity 43.8%; Pred. No. 3.5e-46;  
Matches 183; Conservative 53; Mismatches 93; Indels 89; Gaps 5;  
  
QY 30 VKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLLKYDTTQGRFDGTVEVKE 89  
Db 8 IRVAINGFGRIGRLVFRALLSQKNIIEIVAVNDLTHPDTLAHLLKYDSAHGFEF----- 59  
  
QY 90 GGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAEKD 149  
Db 60 -----KKKVAKDNTLMIDKKVLVFSEKD 84  
  
QY 150 PEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLEAT 209  
Db 85 -----PANLPWAEHNIDIVVEST 102  
  
QY 210 SFFAKKEAAEKHLHANGAKKVVITAPG-GNDVKTVVFNTHDILDGTETVISGASCTTNC 268  
Db 103 GRFVSEEGASLHLQA-GAKRVIISAPAKQKTIKTVVYVNVNHHKIINAEDKIISAASCTTNC 161  
  
QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDMILDPHRRGGDLRRARAGANIVPNSTGAAK 328  
Db 162 LAPMVHVLEKNFGILHGTMTVTTHAYTADQRLQDAPH--SDLRRARAAACNIVPTTTGA 219  
  
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVLDKNVSVDEINAAKKAASNDSFGYTED 388  
Db 220 AIGLVVPEATGKNGMALRVPVLTGSIVELCVALEKDATVEQINQAMKKAASAFRYCED 279  
  
QY 389 PIVSSDIVGSYGSLSFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKI 446  
Db 280 EIVSSDIVGSEHSIFDSKLTNIIIEVDGNKLYKYAWYDNESYVNLVVRVNYCAKL 337  
  
RESULT 13  
G3P\_MYCGE  
ID G3P\_MYCGE STANDARD; PRT; 337 AA.  
AC P47543;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).  
GN GAPA OR GAP OR MG301.  
OS Mycoplasma genitalium.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=96026346; PubMed=7569993;  
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,  
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
RT "The minimal gene complement of Mycoplasma genitalium."  
RL Science 270:397-403(1995).  
RN [2]  
RP SEQUENCE OF 1-81 AND 279-337 FROM N.A.  
RC STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=94075230; PubMed=8253680;  
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;  
RT "A survey of the Mycoplasma genitalium genome by using random  
RT sequencing."

RL J. Bacteriol. 175:7918-7930(1993).  
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
CC -!- PATHWAY: Second phase of glycolysis; first step.  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
CC dehydrogenase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U39710; AAC71523.1; --  
DR EMBL; U02213; AAD12507.1; --  
DR EMBL; U02178; AAD12463.1; --  
DR PIR; C64233; C64233.  
DR HSSP; P17721; IHDG.  
DR TIGR; MG301; --  
DR InterPro; IPR000173; GAP\_dhhydrogenase.  
DR InterPro; IPR006424; GAPDH-I.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh\_C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.  
FT BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE  
FT (BY SIMILARITY).  
FT ACT\_SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS  
FT (BY SIMILARITY).  
SQ SEQUENCE 337 AA; 37097 MW; FA1EAL966687006B CRC64;  
  
Query Match 37.5%; Score 854.5; DB 1; Length 337;  
Best Local Similarity 42.8%; Pred. No. 3.2e-45;  
Matches 179; Conservative 52; Mismatches 98; Indels 89; Gaps 5;  
  
QY 30 VKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLLKYDTTQGRFDGTVEVKE 89  
Db 8 IKVAINGFGRIGRLVFRSLLSKANVEVVAINDLTQPEVLAHLLKYDSAHGELKRKITVKQ 67  
  
QY 90 GGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAEKD 149  
Db 68 NILQIDRKVVVFSEKDPQNLPW-----D 91  
  
QY 150 PEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLEAT 209  
Db 92 EHDID-----VVIEST 102  
  
QY 210 SFFAKKEAAEKHLHANGAKKVVITAPG-GNDVKTVVFNTHDILDGTETVISGASCTTNC 268  
Db 103 GRFVSEEGASLHLKA-GAKRVIISAPAKEKTIRTVVYVNVNHHKTISSDDKIISAASCTTNC 161  
  
QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDMILDPHRRGGDLRRARAGANIVPNSTGAAK 328  
Db 162 LAPLVHVLEKNFGIVGTMTLVTHAYTADQRLQDAPH--NDLRRARAAAVNIVPTTTGA 219  
  
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVLDKNVSVDEINAAKKAASNDSFGYTED 388  
Db 220 AIGLVVPEANGKNGMSLRVPVLTGSIVELSVLEKSPSVEQVQNAKRFASAFKYCED 279  
  
QY 389 PIVSSDIVGSYGSLSFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKI 446  
Db 280 PIVSSDVVSSEYGSIFDSKLTNIVEVDGMKLYKYAWYDNESYVHQLVRVVSYCAKL 337  
  
RESULT 14  
G3P\_CORGL  
ID G3P\_CORGL STANDARD; PRT; 334 AA.

AC Q01651;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).  
GN GAP OR CGL1588.  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=1718;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 13059 / AS019;  
RX MEDLINE=93015645; PubMed=1400158;  
RA Eikmanns B.J.;  
RT "Identification, sequence analysis, and expression of a  
RT Corynebacterium glutamicum gene cluster encoding the three glycolytic  
RT enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate  
RT kinase, and triosephosphate isomerase.";  
RL J. Bacteriol. 174:6076-6086(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
RA Nakagawa S.;  
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
CC -!- PATHWAY: Second phase of glycolysis; first step.  
CC -!- SUBUNIT: Homotetramer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
CC dehydrogenase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X59403; CAA42045.1; -.  
DR EMBL; AP005279; BAB98981.1; -.  
DR PIR; A43260; A43260.  
DR HSP; P00362; 1GD1.  
DR InterPro; IPR000173; GAP\_dhhydrogenase.  
DR InterPro; IPR006424; GAPDH-I.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh.C; 1.  
DR PRINTS; PR00078; G3PDHGRGNASE.  
DR TIGRfams; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.  
FT BINDING 153 153 GLYCERALDEHYDE 3-PHOSPHATE.  
FT ACT SITE 180 180 ACTIVATES THIOL GROUP DURING CATALYSIS.  
FT CONFLICT 25 26 SD -> NG (IN REF. 1).  
FT CONFLICT 333 334 KL -> QALN (IN REF. 1).  
SQ SEQUENCE 334 AA; 36045 MW; 33792AF65FA90FA7 CRC64;  
  
Query Match 37.3%; Score 850; DB 1; Length 334;  
Best Local Similarity 44.7%; Pred. No. 6e-45;  
Matches 189; Conservative 41; Mismatches 95; Indels 98; Gaps 7;  
  
QY 28 MVVKVGINGFRIGRLAFRI-QNVEGVEVTRINDLTDNMLAHLLKYDTTQGRFDGIVE 86  
Db 1 MTRVGVINGFRIGRNFRAVLESDLEAVAVNDLTNKTLLTKFDSIMGRLGQVE 60  
  
QY 87 VKEGGFVNGNFVKVSAERDPENIDWATDGVIEIVLEALEGTVEVKDGGFDVNGKFIKVS 146  
Db 61 YDDDSITVGGKRIAVYAEKDPK----- 82  
  
QY 147 EKDPQIDWATDGVIEIVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEIVL 206

Db 83 -----NLDWAAHNVDIVI 95  
QY 207 EATSFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILD-GTETVISGASCT 265  
Db 96 ESTGFFTDANAAKAHIEA-GAKKVIISAPASNEDATFVYGVNHESDPENHNVISGASCT 154  
QY 266 TNCLAPMAKALHDAFGIQKGLMTTIHAYTGDQMIIDGPHRGDLRRARAGAANIVPNSTG 325  
Db 155 TNCLAPMAKVLNDKFGIENGLMTTVHAYTGDQRLHDAPHR--DLRRARAAAANIVPTSTG 212  
QY 326 AAKAIGLVIPELNGKLDGAAQRPVPTGVSVELVVTLDKNVSVDEINAAKKAASNDSFG- 384  
Db 213 AAKAVALVLPKLGKLDGYALRVPVITGSATDLTFNTKSEVTVESINAAIKEAAVGEFGE 272  
QY 385 ---YTEDPIVSSDIVGSYGLSLEFDTQTVMVEVDGSQLVKVSVWYDNEMSYTAQLVRTLE 441  
Db 273 TLAYSEEPVSTDIVHDSHGSIFDAGLTKV---SGNTVKVSVWYDNEMGYTCQLLRLE 328  
QY 442 YFA 444  
Db 329 LVA 331  
  
RESULT 15  
G3P1\_BACSU  
ID G3P1\_BACSU STANDARD; PRT; 334 AA.  
AC P09124;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH) (NAD-  
DE dependent glyceraldehyde-3-phosphate dehydrogenase).  
GN GAPA OR GAP OR BSU33940.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / BD170;  
RX MEDLINE=89160255; PubMed=2493629;  
RA Viaene A., Dhaese P.;  
RT "Sequence of the glyceraldehyde-3-phosphate dehydrogenase gene from  
RT Bacillus subtilis.";  
RL Nucleic Acids Res. 17:1251-1251(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer J., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche M., Rose M., Sadaie Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takamaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,



RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,  
RA Winters P., Wipatt A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis";  
RL Nature 390:249-256(1997).  
RN [3]  
RP SEQUENCE OF 1-30.  
RC STRAIN=168 / JH642;  
RX MEDLINE=96345629; PubMed=8755892;  
RA Grammann P., Schroeder K., Schmid R., Marahiel M.A.;  
RT "Cold shock stress-induced proteins in Bacillus subtilis.";  
RL J. Bacteriol. 178:4611-4619(1996).  
RN [4]  
RP CHARACTERIZATION.  
RX MEDLINE=20261518; PubMed=10799476;  
RA Fillingner S., Boschi-Muller S., Azza S., Dervyn E., Brantlant G.,  
RA Aymerich S.;  
RT "Two glyceraldehyde-3-phosphate dehydrogenases with opposite  
RT physiological roles in a nonphotosynthetic bacterium.";  
RL J. Biol. Chem. 275:14031-14037(2000).  
CC -!- FUNCTION: More active in catabolism.  
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
CC -!- PATHWAY: Second phase of glycolysis; first step.  
CC -!- SUBUNIT: Homotetramer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
CC dehydrogenase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X13011; CA31434.1; -.  
DR EMBL; Z99121; CAB15399.1; -.  
DR PIR; S02754; DEBSG.  
DR HSSP; P00362; IGD1.  
DR Subtilist; BG10827; gapA.  
DR InterPro; IPR000173; GAP\_dhhydrogenase.  
DR InterPro; IPR006424; GAPDH-I.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.  
FT INIT MET 0 0  
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE.  
FT ACT SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS.  
SQ SEQUENCE 334 AA; 35701 MW; 1283D3E6CF5095EC CRC64;  
  
Query Match 37.1%; Score 844; DB 1; Length 334;  
Best Local Similarity 43.7%; Pred. No. 1.4e-44;  
Matches 184; Conservative 45; Mismatches 96; Indels 96; Gaps 7;  
  
Qy 30 VKVGINGFGRIGRLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEVKE 89  
Db 2 VKVGINGFGRIGRNVFRAALNNPEVAVVNDLTDANMLAHLQYDSVHGKLDAEVSV-- 59  
  
Qy 90 GGFVNGNFIKVSARDPENIDWATDGVIEIVLEALEGTVEVKDGGFDVNGKFIKVSAREKD 149  
Db 60 ----- 59  
  
Qy 150 PEQIDWATDGVIEIVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIWDATDGVIEVLEAT 209  
Db 60 -----DGNLV-----VNGKTIEVSAERDPAKLSWGKQGVIEIVVEST 96

QY 210 SFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVPVNTNHDILD-GTETVISGASCTTNC 268  
Db 97 GFFTkradaAKHLEA-GAKKVIISAPANEEDITIVMGVNEDKYDAANHDVISNASCTTNC 155  
QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDDLRRARAGAAIVPNSTGAAK 328  
Db 156 LAPFAKVLNDKFGIKRGMMTTVHSYTNQQILDLPKH--DYRRARAAAENIPTSTGAAK 213  
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVSVELVVTLDKNVSVDEINAAKASNDS----FG 384  
Db 214 AVSLVLPKELKGLNGGAMRVPTPNVSLVDLVAELNQEVTAEEVNAALKEAAEGDLKGILG 273  
QY 385 YTEDPIVSSDIVGSYGSLSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFA 444  
Db 274 YSEEPVSGDYNNGNKNSSITDALSTMVME---GSMVKVISWYDNESGYSNRVVDLAAYIA 330  
QY 445 K 445  
Db 331 K 331

Search completed: September 15, 2004, 10:24:57  
Job time : 27 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:18:54 ; Search time 119 Seconds  
(without alignments)  
1187.833 Million cell updates/sec

Title: US-10-650-369-22  
Perfect score: 2278  
Sequence: 1 MKKITGIILLALLAVILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000  
Listing first 45 summaries

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	1656.5	72.7 336 2 Q83ZF7	Q83zf7 streptococc
2	1564.5	68.7 336 2 Q9ALW2	Q9alw2 streptococc
3	1564.5	68.7 336 16 Q8E3E8	Q8e3e8 streptococc
4	1564.5	68.7 336 16 Q8DXS8	Q8dxs8 streptococc
5	1559.5	68.5 336 2 Q8KVU6	Q8kvv6 streptococc
6	1544.5	67.8 336 2 Q8GCR7	Q8gcr7 streptococc
7	1535	67.4 335 16 Q97NLI1	Q97nl1 streptococc
8	1535	67.4 359 16 Q8CWN6	Q8cwn6 streptococc
9	1534.5	67.4 336 2 Q8KHG1	Q8khg1 streptococc
10	1530.5	67.2 336 2 Q8KVU5	Q8kvu5 streptococc
11	1521	66.8 337 16 Q8DVV3	Q8dvv3 streptococc
12	1513.5	66.4 336 2 Q8VVB9	Q8vvb9 streptococc
13	1459.5	64.1 320 2 Q9L5X6	Q9l5x6 streptococc
14	1410	61.9 309 2 Q9AJT7	Q9ajt7 streptococc
15	1409	61.9 309 2 Q9AJT9	Q9ajt9 streptococc
16	1407	61.8 309 2 Q9AJT4	Q9ajt4 streptococc

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

17	1405	61.7	309	2	Q9AJT5	Q9ajt5 streptococc
18	1404	61.6	308	2	Q9AJT8	Q9ajt8 streptococc
19	1387	60.9	305	2	Q9AJT6	Q9ajt6 streptococc
20	1357.5	59.6	336	16	Q9CDH4	Q9cdh4 lactococcus
21	1282	56.3	333	16	Q833I8	Q833i8 enterococcu
22	1205.5	52.9	336	16	Q8Y4I1	Q8y4i1 listeria m
23	1204.5	52.9	336	16	Q928H9	Q928h9 listeria in
24	1175	51.6	332	16	Q8XKT9	Q8xkt9 clostridium
25	1164	51.1	335	16	Q8RFN9	Q8rfn9 fusobacteri
26	1145	50.3	334	16	Q9JW78	Q9jw78 neisseria m
27	1138.5	50.0	336	16	Q8CPY5	Q8cpy5 staphylococ
28	1134	49.8	330	2	Q83UX3	Q83ux3 neisseria g
29	1132.5	49.7	336	16	Q9Z5C5	Q9z5c5 staphylococ
30	1132	49.7	334	16	Q9JX95	Q9jx95 neisseria m
31	1131	49.6	330	2	Q84HZ6	Q84hz6 neisseria g
32	1130	49.6	330	2	Q84HZ4	Q84hz4 neisseria g
33	1129	49.6	335	2	Q93M61	Q93m61 staphylococ
34	1127	49.5	330	2	Q84HZ5	Q84hz5 neisseria g
35	1089	47.8	310	2	Q8L2P7	Q8l2p7 staphylococ
36	1082.5	47.5	311	2	Q8L2R0	Q8l2r0 staphylococ
37	1075	47.2	310	2	Q8L2Q2	Q8l2q2 staphylococ
38	1074	47.1	310	2	Q8L2Q6	Q8l2q6 staphylococ
39	1073	47.1	310	2	Q8L2P6	Q8l2p6 staphylococ
40	1071	47.0	310	2	Q8L2P5	Q8l2p5 staphylococ
41	1070	47.0	311	2	Q8VM79	Q8vm79 rhodospiril
42	1069	46.9	310	2	Q8L2Q3	Q8l2q3 staphylococ
43	1068.5	46.9	311	2	Q8L2P8	Q8l2p8 staphylococ
44	1067	46.8	310	2	Q8L2Q7	Q8l2q7 staphylococ
45	1065	46.8	310	2	Q8L2R7	Q8l2r7 staphylococ

ALIGNMENTS

RESULT 1

Q83ZF7	ID	Q83ZF7	PRELIMINARY;	PRT;	336 AA.
AC	Q83ZF7;				
DT	01-JUN-2003	(TremBLrel. 24, Created)			
DT	01-JUN-2003	(TremBLrel. 24, Last sequence update)			
DT	01-OCT-2003	(TremBLrel. 25, Last annotation update)			
DE	Glyceraldehyde 3-P dehydrogenase.				
GN	GAPC.				
OS	Streptococcus dysgalactiae.				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1334;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Perez-Casal J.F., Bolton A., Song X.-M., Willson P., Potter A.A.;				
RT	"Use of the surface proteins GapC and Mig of Streptococcus				
RT	dysgalactiae as protective antigens against mastitis in non-lactating				
RT	cows."				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF375662; AAP31408.1; -.				
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.				
DR	GO; GO:0006096; P:glycolysis; IEA.				
DR	InterPro; IPR000173; GAPDH-I.				
DR	InterPro; IPR000173; GAP_dhydrogenase.				
DR	Pfam; PF00044; gpdh; 1.				
DR	Pfam; PF02800; gpdh_C; 1.				
DR	PRINTS; PR00078; G3PDHGRGNASE.				
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.				
DR	PROSITE; PS00071; GAPDH; 1.				
SQ	SEQUENCE 336 AA; 35928 MW; 11828218CF037076 CRC64;				

Query Match 72.7%; Score 1656.5; DB 2; Length 336;  
Best Local Similarity 79.6%; Pred. No. 2.3e-88;  
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

Qy	28	MVVKVGINGFGRIGRLAFRIQNVGVETRIINDLTDPNMLAHLKDYDTTQGRFDGTVEV	87
Db	1	MVVKVGINGFGRIGRLAFRIQNVGVETRIINDLTDPNMLAHLKDYDTTQGRFDGTVEV	60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAB 147  
Db 61 KEGGFEVNGNFIKVSARDPE-----81  
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 207  
Db 82 -----NIDWATDGVLEIVLE 95  
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVVENTNHDILDGTETVISGASCTTN 267  
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVVENTNHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 327  
Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSFGYTE 275  
QY 388 DPIVSSDIVGSYGSLEFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGSYGSLEFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336

RESULT 2  
Q9ALW2 Q9ALW2 PRELIMINARY; PRT; 336 AA.  
AC Q9ALW2;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase.  
OS Streptococcus agalactiae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1311;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=J48;  
RA Seifert K.N., Bleiweis A.S., McArthur W.P., Brady L.J.;  
RT "The Group B Streptococcal Surface Antigen Delta is a Glyceraldehyde 3-Phosphate Dehydrogenase."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.  
DR EMBL; AF338416; AAK14387.1; -.  
DR HSSP; P00362; 1GDI.  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR006424; GAPDH-I.  
DR InterPro; IPR000173; GAP\_dhdrogenase.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Oxidoreductase.  
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 68.7%; Score 1564.5; DB 2; Length 336;  
Best Local Similarity 74.1%; Pred. No. 5e-83;  
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;  
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRF-----54

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAB 147  
Db 55 -----54  
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 207  
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 95  
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVVENTNHDILDGTETVISGASCTTN 267  
Db 96 ATGFFAKKEAAEQHTIHENGAKKVVITAPGGNDVKTVVVENTNHDILDGTETVISGASCTTN 155  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 215  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSFGYTE 275  
QY 388 DPIVSSDIVGSYGSLEFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGSYGSLEFDATQTKVQTVDGNQLVKVSVWYDNEMSYTSQLVRTLEYFAKIA 335  
QY 448 K 448  
Db 336 K 336

RESULT 3  
Q8E3E8 Q8E3E8 PRELIMINARY; PRT; 336 AA.  
AC Q8E3E8;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase.  
GN GBS1811.  
OS Streptococcus agalactiae (serotype III).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=216495;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NEM316 / Serotype III;  
RX MEDLINE=22242508; PubMed=12354221;  
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L., Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P., Kunst F.;  
RA "Genome sequence of Streptococcus agalactiae, a pathogen causing RT invasive neonatal disease."  
RT Mol. Microbiol. 45:1499-1513 (2002).  
RL EMBL; AL766853; CAD47470.1; -.  
DR Sagalists; gbs1811; -.  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR006424; GAPDH-I.  
DR InterPro; IPR000173; GAP\_dhdrogenase.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Complete proteome.  
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 68.7%; Score 1564.5; DB 16; Length 336;  
Best Local Similarity 74.1%; Pred. No. 5e-83;  
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;  
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRF-----54

Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRF----- 54

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147

Db 55 ----- 54

QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207

Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 95

QY 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267

Db 96 ATGFFASKEKAHQHENGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAPFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 327

Db 156 CLAPMAKALQDNFVKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKNVSVDEINAAKKAASNDSEGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVATLEKDVTVVEEVNAAKKAANDSYGYTE 275

QY 388 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447

Db 276 DPIVSSDIVGSYSLFDATQTKVQTVTDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 4

Q8DXS8 PRELIMINARY; PRT; 336 AA.

AC Q8DXS8;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Glyceraldehyde 3-phosphate dehydrogenase.

GN GAP OR SAG1768.

OS Streptococcus agalactiae (serotype V).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=216466;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=2603 V/R / Serotype V;

RX MEDLINE=2222988; PubMed=12200547;

RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S., Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D., Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R., Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S., Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M., Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D., Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G., Fraser C.M.;

RA "Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).

DR EMBL; AE014272; AAN00631.1; --

DR TIGR; SAG1768; --

DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.

DR GO; GO:0006096; P:glycolysis; IEA.

DR InterPro; IPR006424; GAPDH-I.

DR InterPro; IPR000173; GAP\_dhhydrogenase.

DR Pfam; PF00044; gpdh; 1.

DR Pfam; PF02800; gpdh C; 1.

DR PRINTS; PR00078; G3PDHDRGNASE.

DR TIGRFAMs; TIGR01534; GAPDH-I; 1.

DR PROSITE; PS00071; GAPDH; 1.

KW Complete proteome.

SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 68.7%; Score 1564.5; DB 16; Length 336;

Best Local Similarity 74.1%; Pred. No. 5e-83;

Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87

Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRF----- 54

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147

Db 55 ----- 54

QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207

Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 95

QY 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267

Db 96 ATGFFASKEKAHQHENGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAPFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 327

Db 156 CLAPMAKALQDNFVKQGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKNVSVDEINAAKKAASNDSEGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVATLEKDVTVVEEVNAAKKAANDSYGYTE 275

QY 388 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447

Db 276 DPIVSSDIVGSYSLFDATQTKVQTVTDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 5

Q8KVU6 PRELIMINARY; PRT; 336 AA.

AC Q8KVU6;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Glyceraldehyde-3-phosphate dehydrogenase.

OS Streptococcus uberis.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=1349;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 9927;

RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J., Potter A.A.;

RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP antigens confers protection against heterologous challenge with Streptococcus uberis.";

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.

CC EMBL; AF421900; AAM73771.1; --

DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006096; P:glycolysis; IEA.

DR InterPro; IPR006424; GAPDH-I.

DR InterPro; IPR000173; GAP\_dhhydrogenase.

DR Pfam; PF00044; gpdh; 1.

DR Pfam; PF02800; gpdh C; 1.

DR PRINTS; PR00078; G3PDHDRGNASE.

DR TIGRFAMs; TIGR01534; GAPDH-I; 1.

DR PROSITE; PS00071; GAPDH; 1.

KW Oxidoreductase.



SQ	SEQUENCE	336 AA;	35906 MW;	755D74A4548E61D4	CRC64;
Query Match					
Best Local Similarity 68.5%; Score 1559.5; DB 2; Length 336;					
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;					
QY	28	MVVKVGINGFGRIGRLAFRRRIQNVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV	87		
Db	1	MVVKVGINGFGRIGRLAFRRRIQNVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV	60		
QY	88	KEGGFEVNGNFIKVSAAERDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVS	147		
Db	61	KDGGFEVNGNFIKVSAAEKDPE	81		
QY	148	KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE	207		
Db	82	-----	95		
QY	208	ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	267		
Db	96	ATGFFAKKAAAEKHLHANGAKKVVITAPGGDDVKTVVFNTHDILDGTETVISGASCTTN	155		
QY	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA	327		
Db	156	CLAPMAKALQDNFVKQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGASNIVPNSTGAA	215		
QY	328	KAIGLVIPELNGKLDGAAQRPVPTGSVTELVTLDKNVSVDEINAAKAAANDSFGYTE	387		
Db	216	KAIGLVIPELNGKLDGAAQRPVPTGSVTELVAVLEKETSVEEINAAKAAANDSYGYTE	275		
QY	388	DPIVSSDIVGSYGSLEFDTQTQKMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA	447		
Db	276	DPIVSSDIIIGMAYGSLFDTQTQKVQTVDGNQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA	335		
QY	448	K 448			
Db	336	K 336			
RESULT 6					
Q8GCR7		PRELIMINARY;	PRT;	336 AA.	
ID	Q8GCR7				
AC	Q8GCR7;				
DT	01-MAR-2003	(TrEMBLrel. 23, Created)			
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)			
DE	Glyceraldehyde-3-phosphate	dehydrogenase (EC 1.2.1.12).			
OS	Streptococcus suis.				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1307;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=S735;				
RA	Brassard J., Gottschalk M., Quessy S.;				
RT	"Cloning and purification of Streptococcus suis serotype 2				
RT	glyceraldehyde-3-phosphate dehydrogenase.";				
RL	Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY167026; AAN86058.1; -.				
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.				
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.				
DR	GO; GO:0006096; P:glycolysis; IEA.				
DR	InterPro; IPR006424; GAPDH-I.				
DR	InterPro; IPR00173; GAP_dhhydrogenase.				
DR	Pfam; PF00044; gpdh; 1.				
DR	Pfam; PF02800; gpdh C; 1.				
DR	PRINTS; PR00078; G3PDHDRGNASE.				
DR	TIGRFAMs; TIGR01534; GAPDH-I; 1.				
DR	PROSITE; PS00071; GAPDH; 1.				
KW	Oxidoreductase.				
SQ	SEQUENCE	336 AA;	35825 MW;	1DEB9B1A492DCF59	CRC64;
Query Match					
67.8%; Score 1544.5; DB 2; Length 336;					

Best Local Similarity 72.9%; Pred. No. 7.2e-82;					
Matches 307; Conservative 16; Mismatches 13; Indels 85; Gaps 1;					
QY	28	MVVKVGINGFGRIGRLAFRRRIQNVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV	87		
Db	1	MVVKVGINGFGRIGRLAFRRRIQNVEGVETRINDLTDPNMLAHLKYDTTQGRF	54		
QY	88	KEGGFEVNGNFIKVSAAERDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVS	147		
Db	55	-----	54		
QY	148	KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE	207		
Db	55	-----DGTVEVKDGGFEVNGKFKVSAEREPGNIDWATDGVDIVLE	95		
QY	208	ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	267		
Db	96	ATGFFASKEKAEQHIHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	155		
QY	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHRGGDLRRARAGAAANIVPNSTGAA	327		
Db	156	CLAPMAKALHDAFGVQKGLMTTIHGTYTGDQWLDGHRGGDLRRARAAAAANIVPNSTGAA	215		
QY	328	KAIGLVIPELNGKLDGAAQRPVPTGSVTELVTLDKNVSVDEINAAKAAANDSFGYTE	387		
Db	216	KAIGLVIPELNGKLDGAAQRPVPTGSVTELVATLDKKVTAEEVNAAMKAAATESFGYTE	275		
QY	388	DPIVSSDIVGSYGSLEFDTQTQKMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA	447		
Db	276	DQLVSSDIVGISFGSLFDTQTQKVIEWDGEQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA	335		
QY	448	K 448			
Db	336	K 336			

RESULT 7

Q97NLI		PRELIMINARY;	PRT;	335 AA.
ID	Q97NLI			
AC	Q97NLI;			
DT	01-OCT-2001	(TrEMBLrel. 18, Created)		
DT	01-OCT-2001	(TrEMBLrel. 18, Last sequence update)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)		
DE	Glyceraldehyde 3-phosphate	dehydrogenase (EC 1.2.1.12).		
GN	SP2012 OR GAPDH.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1313;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC BAA-334 / TIGR4;			
RX	MEDLINE=21357209; PubMed=11463916;			
RA	Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,			
RA	Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,			
RA	Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,			
RA	Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,			
RA	Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,			
RA	McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,			
RA	Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,			
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;			
RT	"Complete genome sequence of a virulent isolate of Streptococcus			
RT	pneumoniae.";			
RL	Science 293:498-506(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC11733;			
RA	Bergmann S., Hammerschmidt S.;			
RT	"Identification of pneumococcal GAPDH as plasmin(ogen)-binding			
RT	protein.";			
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AE007490; AAK76079.1; -.			
DR	EMBL; AJ505822; CAD44376.1; -.			



DR PIR; F95235; F95235.  
DR HSSP; P00354; 3GPD.  
DR TIGR; SP2012; -.  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR006424; GAPDH-I.  
DR InterPro; IPR000173; GAP\_dhdrogenase.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Oxidoreductase; Complete proteome.  
SQ SEQUENCE 335 AA; 35856 MW; DA483CEA423E747B CRC64;

Query Match 67.4%; Score 1535; DB 16; Length 335;  
Best Local Similarity 73.6%; Pred. No. 2.5e-81;  
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVINGFGGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGGRIGRLAFRRIONVEGVETRLNDLTPVMLAHLKYDTTQGRFDGTVEV 60  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEIDGTVEVKDGGFDVNGKFIKVSAAE 147  
Db 61 KEGGFEVNGKFIKVSARDPE----- 81  
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 207  
Db 82 -----QIDWATDGVLEIVLE 95  
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDTGTETVISGASCTTN 267  
Db 96 ATGFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVVNTNHDVLDGTETVISGASCTTN 154  
QY 268 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSTGAA 327  
Db 155 CLAPMAKALQDNFVVEGLMTTIHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSTGAA 214  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTTELVTLDKNVSVDEINAAKKAASNDSFGYTE 387  
Db 215 KAIGLVIPELNGKLDGSAQRVPTGVTTELVAVLEKNVTDEVNAAKKAASNESYGYTE 274  
QY 388 DPIVSSDIVGVSYSGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 275 DPIVSSDIVGMSYSGLFDATQTKVLDVDGKQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 334  
QY 448 K 448  
Db 335 K 335

RESULT 8  
Q8CWN6 PRELIMINARY; PRT; 359 AA.  
ID Q8CWN6  
AC Q8CWN6;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase (Phosphorylating)  
DE (EC 1.2.1.12).  
GN GAP A OR SPR1825.  
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=171101;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21429245; PubMed=11544234;  
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,  
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,  
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,  
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,

RA McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I.,  
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,  
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,  
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rosteck P.R. Jr., Skatrud P.L.,  
RA Glass J.I.;  
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";  
RL J. Bacteriol. 183:5709-5717(2001).  
DR EMBL; AE008547; AAL00628.1; -.  
DR PIR; G98099; G98099.  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR006424; GAPDH-I.  
DR InterPro; IPR000173; GAP\_dhdrogenase.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Oxidoreductase; Complete proteome.  
SQ SEQUENCE 359 AA; 38763 MW; 61EF0E375B330B60 CRC64;

Query Match 67.4%; Score 1535; DB 16; Length 359;  
Best Local Similarity 73.6%; Pred. No. 2.8e-81;  
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVINGFGGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87  
Db 25 MVVKVINGFGGRIGRLAFRRIONVEGVETRLNDLTPVMLAHLKYDTTQGRFDGTVEV 84  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEIDGTVEVKDGGFDVNGKFIKVSAAE 147  
Db 85 KEGGFEVNGKFIKVSARDPE----- 105  
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 207  
Db 106 -----QIDWATDGVLEIVLE 119  
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDTGTETVISGASCTTN 267  
Db 120 ATGFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVVNTNHDVLDGTETVISGASCTTN 178  
QY 268 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSTGAA 327  
Db 179 CLAPMAKALQDNFVVEGLMTTIHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSTGAA 238  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTTELVTLDKNVSVDEINAAKKAASNDSFGYTE 387  
Db 239 KAIGLVIPELNGKLDGSAQRVPTGVTTELVAVLEKNVTDEVNAAKKAASNESYGYTE 298  
QY 388 DPIVSSDIVGVSYSGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 299 DPIVSSDIVGMSYSGLFDATQTKVLDVDGKQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 358  
QY 448 K 448  
Db 359 K 359

RESULT 9  
Q8KHG1 PRELIMINARY; PRT; 336 AA.  
ID Q8KHG1  
AC Q8KHG1;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glyceraldehyde-3-phosphate dehydrogenase.  
OS Streptococcus agalactiae, and  
OS Streptococcus iniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1311, 1346;  
RN [1]

RP SEQUENCE FROM N.A.  
RC SPECIES=S.agalactiae, and S.iniae; STRAIN=ATCC 27541, and 9117;  
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,  
RA Potter A.A.;  
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP  
RT antigens confers protection against heterologous challenge with  
RT Streptococcus uberis.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE  
CC DEHYDROGENASE FAMILY.  
DR EMBL; AF421899; AAM73770.1; --  
DR EMBL; AF421902; AAM73773.1; --  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR006424; GAPDH-I.  
DR InterPro; IPR00173; GAP\_dhdrogenase.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Oxidoreductase.  
SQ SEQUENCE 336 AA; 35723 MW; ABAB9E14F3ED1111 CRC64;

Query Match 67.4%; Score 1534.5; DB 2; Length 336;  
Best Local Similarity 72.7%; Pred. No. 2.7e-81;  
Matches 306; Conservative 17; Mismatches 13; Indels 85; Gaps 1;  
  
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 60  
  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
Db 61 KDGGEVNGS----- 70  
  
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 207  
Db 71 -----FVKVSAEREPANIDWATDGVLEIVLE 95  
  
QY 208 ATSFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFASKAAAEQIHANGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155  
  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHGVTGQMVLDDGPHRGDLRRARAAAAANIVPNSTGAA 215  
  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKAAASNDSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVAVLEKDTSVLEINAAKAAANDSYGYTE 275  
  
QY 388 DPIVSSDIVGSYGSLEFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DAIVSSDIVGISYGSLEFDATQTKVQTVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335  
  
QY 448 K 448  
Db 336 K 336

RESULT 10  
Q8KVU5  
ID Q8KVU5 PRELIMINARY; PRT; 336 AA.  
AC Q8KVU5;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glyceraldehyde-3-phosphate dehydrogenase.  
OS Streptococcus parauberis.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.

OX NCBI\_TaxID=1348;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,  
RA Potter A.A.;  
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP  
RT antigens confers protection against heterologous challenge with  
RT Streptococcus uberis.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE  
CC DEHYDROGENASE FAMILY.  
DR EMBL; AF421901; AAM73772.1; --  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR006424; GAPDH-I.  
DR InterPro; IPR00173; GAP\_dhdrogenase.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Oxidoreductase.  
SQ SEQUENCE 336 AA; 36037 MW; 311647C25489AC9E CRC64;

Query Match 67.2%; Score 1530.5; DB 2; Length 336;  
Best Local Similarity 73.4%; Pred. No. 4.6e-81;  
Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;  
  
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 55  
  
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
Db 56 -----GTVEVKDGGFDVNGKFIKVSAAE 77  
  
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 207  
Db 78 KDPEQ-----IDWATDGVLEIVLE 95  
  
QY 208 ATSFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267  
Db 96 ATGFFAKKAAAEKHLHENGAKKVITAPGGDDVKTVVFNTHDILDGTETVISGASCTTN 155  
  
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDMILLDGPGRGDLRRARAGANNIVPNSTGAA 215  
  
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKAAASNDSFGYTE 387  
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVAVLNKETSVEEINSVMKAAANDSYGYTE 275  
  
QY 388 DPIVSSDIVGSYGSLEFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGMSFGSLEFDATQTKVQTVDGNQLVKVSWYDNEMSYTAQLDRTLEYFAKIA 335  
  
QY 448 K 448  
Db 336 K 336

RESULT 11  
Q8DVV3  
ID Q8DVV3 PRELIMINARY; PRT; 337 AA.  
AC Q8DVV3;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Extracellular glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12).  
GN GAPC OR SMU.360.  
OS Streptococcus mutans.

OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	OS	Streptococcus thermophilus.
OC	Streptococcus.	OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX	NCBI_TaxID=1309;	OC	Streptococcus.
RN	[1]	OX	NCBI_TaxID=1308;
RP	SEQUENCE FROM N.A.	RN	[1]
RC	STRAIN=UA159 / ATCC 700610 / Serotype C;	RP	SEQUENCE FROM N.A.
RX	MEDLINE=22295063; PubMed=12397186;	RC	STRAIN=LMG18311;
RA	Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,	RA	van den Bogaard P.T.C., Kleerebezem M., Hols P., Crispie F.,
RA	Carlson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,	RA	Kuipers O.P., de Vos W.M.;
RA	Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;	RT	"Modulation of Glycolysis by Lactose Availability in Streptococcus
RT	"Genome sequence of Streptococcus mutans UA159, a cariogenic dental	RT	thermophilus.";
RT	pathogen.";	RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RL	Proc. Natl. Acad. Sci. U.S.A: 99:14434-14439 (2002).	CC	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
DR	EMBL; AE014883; AAN58118.1; -.	CC	DEHYDROGENASE FAMILY.
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.	DR	EMBL; AF442551; AAL35377.1; -.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.	DR	HSSP; P00354; 3GPD.
DR	GO; GO:0006096; P:glycolysis; IEA.	DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR	InterPro; IPR006424; GAPDH-I.	DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	InterPro; IPR000173; GAP_dhhydrogenase.	DR	GO; GO:0006096; P:glycolysis; IEA.
DR	Pfam; PF00044; gpdh; 1.	DR	InterPro; IPR006424; GAPDH-I.
DR	Pfam; PF02800; gpdh.C; 1.	DR	InterPro; IPR000173; GAP_dhhydrogenase.
DR	PRINTS; PR00078; G3PDHDRGNASE.	DR	Pfam; PF02800; gpdh.C; 1.
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.	DR	PRINTS; PR00078; G3PDHDRGNASE.
DR	PROSITE; PS00071; GAPDH; 1.	DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.
KW	Oxidoreductase; Complete proteome.	DR	PROSITE; PS00071; GAPDH; 1.
SQ	SEQUENCE 337 AA; 36068 MW; 42BFE20365963C22 CRC64;	KW	Oxidoreductase.
		SQ	SEQUENCE 336 AA; 36026 MW; 52C1F25F3A7E0230 CRC64;
	Query Match 66.8%; Score 1521; DB 16; Length 337;		Query Match 66.4%; Score 1513.5; DB 2; Length 336;
	Best Local Similarity 71.8%; Pred. No. 1.7e-80;		Best Local Similarity 71.7%; Pred. No. 4.5e-80;
	Matches 303; Conservative 18; Mismatches 15; Indels 86; Gaps 2;		Matches 302; Conservative 18; Mismatches 16; Indels 85; Gaps 1;
QY	28 MVVKVGINGFGRIGRLAFRRIQNVGEVETRINDLTPNMLAHLKLYDTTQGRFDGTV 87	QY	28 MVVKVGINGFGRIGRLAFRRIQNVGEVETRINDLTPNMLAHLKLYDTTQGRFDGTV 87
Db	1 MVVKVGINGFGRIGRLAFRRIQNVGEVETRINDLTPNMLAHLKLYDSTQGRFDGN 60	Db	1 MVVKVGINGFGRIGRLAFRRIQNVGEVETRINDLTPVMLAHLKLYDTTQGRFDGT 60
QY	88 KEGGFEVNGNFIKVSARDPENIDWATDGV EIVLEALEGTV EVKDGGFDVNGKFIKVS 147	QY	88 KEGGFEVNGNFIKVSARDPENIDWATDGV EIVLEALEGTV EVKDGGFDVNGKFIKVS 147
Db	61 KEGGFEVNGKFKVSAERDPE----- 81	Db	61 KEGGFEVNGKFKVSAERDP E----- 81
QY	148 KDPEQIDWATDGV EIVLEIDGTV EVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207	QY	148 KDPEQIDWATDGV EIVLEIDGTV EVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207
Db	82 -----QIDWAADGVEIVLE 95	Db	82 -----QIDWANDGVEIVLE 95
QY	208 ATSFPAKKEAAEKHLHAN-GAKKVITAPGGNDVKT VVFNTHDILDGTETVISGASCTT 266	QY	208 ATSFPAKKEAAEKHLHANGAKKVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 267
Db	96 ATGFFASKAAAEKHLHANGAKKVITAPGGNDIKTIVFNTHDVLDTETVISGASCTT 155	Db	96 ATGFFTKKVLAEKHLHPGGAKKVITAPGGNDVKTIVFNTHDILDGTETVISGASCTTN 155
QY	267 NCLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGGLRRARAGAAANIVPNSTGA 326	QY	268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGGLRRARAGAAANIVPNSTGAA 327
Db	156 NCLAPMAKALHDFNSIKEGLMTTHAYTGDMVLDGPHRKGLRRARAAAANIVPNSTGA 215	Db	156 CLAPMAKALNDNFGIVEGLMTTHAYTGDMILDGPNRGGLRRARAGAAANIVPNSTGAA 215
QY	327 AKAIGLVIPELNGKLDGAAQRPVPTGTVTELVTLDKNVSVDEINAAKAAASNDSEFGYT 386	QY	328 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVTLDKNVSVDEINAAKAAASNDSEFGYTE 387
Db	216 AKAIGLVIPELNGKLDGAAQRPVPTGTVTELVAVLDDKKVTVDEVNAAKAAANESYGT 275	Db	216 KAIGLVIPELNGKLDGSAQRPVPTGTVTELVAVCEKNVTVDEVNAAKAAATNESYGYTE 275
QY	387 EDPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKI 446	QY	388 DPIVSSDIVGSYGSFLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db	276 EDPIVSSDIVGMSFGSLFDATQTKVLDVDGKQLVKVSWYDNEMSYTSQLVRTLEYFAKI 335	Db	276 DPIVSSDIVGMSYGSFLFDATQTKVLDVDGKQLVKVSWYDNEMSYTSQLVRTLEYFAKIA 335
QY	447 AK 448	QY	448 K 448
Db	336 AK 337	Db	336 K 336
RESULT 12		RESULT 13	
Q8VVB9		Q9L5X6	
ID	Q8VVB9	ID	Q9L5X6
AC	Q8VVB9;	AC	Q9L5X6;
DT	01-MAR-2002 (TrEMBLrel. 20, Created)	DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)	DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE	Glyceraldehyde-3-phosphate dehydrogenase.	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
GN	GAPDH.		



DE Extracellular glyceraldehyde-3-phosphate dehydrogenase  
DE (Fragment).  
OS Streptococcus gordonii.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1302;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Nelson D., Goldstein J.M., Boatright K., Harty D.W.S., Cook S.L.,  
RA Hickman P.J., Potempa J., Travis J., Mayo J.A.;  
RT "Purification and characterization of an extracellular  
RT glyceraldehyde-3-phosphate dehydrogenase from Streptococcus sanguis  
RT and cloning of the gene encoding this enzyme."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE EMBL/GenBank/DBJ databases.  
CC DEHYDROGENASE FAMILY.  
DR EMBL; AF247678; AAF64063.1; -.  
DR HSSP; P00362; 1GD1.  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR006424; GAPDH-I.  
DR InterPro; IPR000173; GAP\_dhrogenase.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh\_C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Oxidoreductase.  
FT NON\_TER 1 1  
FT NON\_TER 320 320  
SQ SEQUENCE 320 AA; 33968 MW; 6FFCC18BB01E91F0 CRC64;

Query Match 64.1%; Score 1459.5; DB 2; Length 320;  
Best Local Similarity 71.9%; Pred. No. 5.7e-77;  
Matches 291; Conservative 14; Mismatches 15; Indels 85; Gaps 1;  
QY 31 KVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKDYDTTQGRFDGTVVEVKEG 90  
Db 1 KVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKDYDTTQGRFDGTVVEVKEG 60  
QY 91 GFEVNGNFIKVSARDPENIDWATDGVIEIVLEALEGTVVEVKDGGFDVNGKFIKVSAEKDP 150  
Db 61 GFEVNGKFIKVSARDPE-----78  
QY 151 EQIDWATDGVIEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEIVLEATS 210  
Db 79 -----NIDWANDGVIEIVLEATG 95  
QY 211 FFAKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLA 270  
Db 96 FFATKAAAEEKHLHAGGAKKVITAPGGSVDKTVVFNTHDILDGTETVISGASCTTNCLA 155  
QY 271 PMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSGTGAAKAI 330  
Db 156 PMAKALQDNFGVVEGLMTTIHAYTGDQMVLDGPHRKGDLLRRARAAAANIVPNSGTGAAKAI 215  
QY 331 GLVIPELNGKLDGAAQRPVPPTGSVTELVVTLTKNSVDEINAAMKAASNDSFGYTEDPI 390  
Db 216 GLVIPELNGKLDGAAQRPVPPTGSVTELVAVLDKNVTVEVNAAMKAASNESYGYTEDPI 275  
QY 391 VSSDIVGVSIGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQ 435  
Db 276 VSSDIVGMAYGSLFDATQTKVLVDVGKQLVKVSWYDNEMSYTAQ 320

RESULT 14  
Q9AJT7  
ID Q9AJT7 PRELIMINARY; PRT; 309 AA.  
AC Q9AJT7;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)  
DE (Fragment).  
GN GAP.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=clinical isolate;  
RA Amezaga M.R., Carter P.E., Cash P., McKenzie H.;  
RT "Sequence variation in mef(A) and gap genes in M-phenotype  
RT erythromycin resistant Streptococcus pneumoniae."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE  
CC DEHYDROGENASE FAMILY.  
DR EMBL; AJ292048; CAC27448.1; -.  
DR HSSP; P00362; 1GD1.  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR006424; GAPDH-I.  
DR InterPro; IPR000173; GAP\_dhrogenase.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh\_C; 1.  
DR PRINTS; PR00078; G3PDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Oxidoreductase.  
FT NON\_TER 1 1  
FT NON\_TER 309 309  
SQ SEQUENCE 309 AA; 32930 MW; 89E45C3BEA6A0528 CRC64;

Query Match 61.9%; Score 1410; DB 2; Length 309;  
Best Local Similarity 71.9%; Pred. No. 4e-74;  
Matches 284; Conservative 12; Mismatches 13; Indels 86; Gaps 2;  
QY 36 GFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKDYDTTQGRFDGTVVEVKEGGFEV 95  
Db 1 GFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKDYDTTQGRFDGTVVEVKEGGFEV 60  
QY 96 GNFIKVSARDPENIDWATDGVIEIVLEALEGTVVEVKDGGFDVNGKFIKVSAEKDPEQIDW 155  
Db 61 GKFIKVSARDPE-----73  
QY 156 ATDGVIEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEIVLEATSFPAKK 215  
Db 74 -----QIDWATDGVIEIVLEATGFFAKK 95  
QY 216 EAAEKHLHANGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKA 275  
Db 96 EAAEKHL-KGGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKA 154  
QY 276 LHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSGTGAAKAIGLVIP 335  
Db 155 LQDNFGVVEGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSGTGAAKAIGLVIP 214  
QY 336 ELNGKLDGAAQRPVPPTGSVTELVVTLTKNSVDEINAAMKAASNDSFGYTEDPIVSSDI 395  
Db 215 ELNGKLDGSAQRPVPTGSVTELVAVLEKNVTVEVNAAMKAASNESYGYTEDPIVSSDI 274  
QY 396 VGSYGSGLFDATQTKVMEVDGSQLVKVSWYDNEM 430  
Db 275 VGSYGSGLFDATQTKVLVDVGKQLVKVSWYDNEM 309

RESULT 15  
Q9AJT9  
ID Q9AJT9 PRELIMINARY; PRT; 309 AA.  
AC Q9AJT9;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)  
DN (Fragment).  
GN GAP.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=clinical isolate;  
RA Amezaga M.R., Carter P.E., Cash P., McKenzie H.;  
RT "Sequence variation in mef(A) and gap genes in M-phenotype  
RT erythromycin resistant Streptococcus pneumoniae.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE  
CC DEHYDROGENASE FAMILY.  
DR EMBL; AJ292046; CAC27446.1; -.  
DR HSSP; P00362; 1GD1.  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR006424; GAPDH-I.  
DR InterPro; IPR001173; GAP\_dhdrogenase.  
DR Pfam; PF00044; gpdh; 1.  
DR Pfam; PF02800; gpdh\_C; 1.  
DR PRINTS; PR00078; G3FDHDRGNASE.  
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW Oxidoreductase.  
FT NON\_TER 1  
FT NON\_TER 309 309  
SQ SEQUENCE 309 AA; 32916 MW; 39E9CC3BEA6F8959 CRC64;

Query Match		61.9%;	Score 1409;	DB 2;	Length 309;
Best Local Similarity		71.6%;	Pred. No. 4.6e-74;		
Matches 283;		Conservative 13;	Mismatches 13;	Indels 86;	Gaps 2;
QY	36	GFGRIQLAFRRIONVEGVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEVKEGGFEVN	95		
Db	1	GFGRIQLAFRRIONVEGVEVTRINDLTDPVMLAHLKDYDTTQGRFDGTVEVKEGGFEVN	60		
QY	96	GNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAEKDPEQIDW	155		
Db	61	GKFVKVSAERDPE-----	73		
QY	156	ATDGVIEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEIVLEATSEFFAKK	215		
Db	74	-----QIDWATDGVIEIVLEATGFFAKK	95		
QY	216	EAAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKA	275		
Db	96	EAAEKHL-KGGAKKVVITAPGNDVKTVVNTNHDVLDGTETVISGASCTTNCLAPMAKA	154		
QY	276	LHDAFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAANIVPNSTGAAKAIGLVIP	335		
Db	155	LQDNFVVEGLMTTIHAYTGDMILDGPHRGDLRRARAGAANIVPNSTGAAKAIGLVIP	214		
QY	336	ELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDINAAMKAASNDSEFGYTEDPVSSDI	395		
Db	215	ELNGKLDGSAQRVPTGVSVELVAVLEKNVTVDVNAAMKAASNESYGYTEDPIVSSDI	274		
QY	396	VGVSYGSLFDATQTKVMEVDGSQLVKVSWYDNEM	430		
Db	275	VGMSYGSLFDATQTKVLDVKGQLVKVSWYDNEM	309		

**this Page Blank (uspto)**